Haag's Theorem
in
Renormalisable Quantum Field Theories

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr.rer.nat.)

im Fach Physik

Spezialisierung Theoretische Physik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin
von

Dipl-Ing. Lutz Klaczynski

Präsident der Humboldt-Universität zu Berlin:

Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:

Prof. Dr. Elmar Kulke

Gutachter:

1. Prof.Dr. Dirk Kreimer
2. Prof.Dr. David Broadhurst
3. Prof.Dr. Raimar Wulkenhaar

Tag der Verteidigung: 6.November 2015

HUMBOLDT-UNIVERSITÄT ZU BERLIN

Abstract

We review a package of triviality results and no-go theorems in axiomatic quantum field theory. Because the concept of operator-valued distributions in this framework comes very close to what we believe canonical quantum fields are about, these results are of consequence to canonical quantum field theory: they suggest the seeming absurdity that this highly victorious theory is incapable of describing interactions.

Of particular interest is Haag's theorem. It essentially says that the unitary intertwiner of the interaction picture does not exist unless it is trivial. We single out unitary equivalence as the most salient provision of Haag's theorem and critique canonical perturbation theory for scalar fields to argue that canonically renormalised quantum field theory bypasses Haag's theorem by violating this very assumption. Since canonical quantum fields are not mathematically well-defined objects, this cannot be proven. We therefore content ourselves with a heuristic argument which we nevertheless deem sufficiently convincing.

We opine that to define a quantum field theory, nonperturbative equations are necessary. The Hopf-algebraic approach to perturbative quantum field theory allows us to derive Dyson-Schwinger equations and the Callan-Symanzik equation in a mathematically sound way, albeit starting with a purely combinatorial setting. We present a pedagogical account of this method and discuss an ordinary differential equation for the anomalous dimension of the photon. A toy model version of this equation can be solved exactly; its solution exhibits an interesting nonperturbative feature whose effect on the running coupling and the self-energy of the photon we investigate. Such nonperturbative contributions may exclude the existence of a Landau pole, an issue that we also discuss.

On the working hypothesis that the anomalous dimension of a quantum field falls into the class of resurgent functions, we study what conditions Dyson-Schwinger and renormalisation group equations impose on its resurgent transseries. We find that under certain conditions, they encode how the perturbative sector determines the nonperturbative one completely.

Zusammenfassung

Wir betrachten eine Reihe von Trivialitätsresultaten und No-Go-Theoremen aus der Axiomatischen Quantenfeldtheorie. Da das Konzept der operatorwertigen Distributionen dem der kanonischen Quantenfelder aus unserer Sicht sehr nahekommt, sind diese Resultate nicht ohne Konsequenz für die kanonische Quantenfeldtheorie: sie legen die scheinbar absurde Behauptung nahe, dass der Formalismus dieser hochgradig erfolgreichen Theorie nicht in der Lage sei, Wechselwirkungen zu beschreiben.

Von besonderem Interesse ist Haags Theorem. Im Wesentlichen sagt es aus, dass der unitäre Intertwiner des Wechselwirkungsbildes nicht existiert oder trivial ist. Als wichtigste Voraussetzung von Haags Theorem arbeiten wir die unitäre Äquivalenz heraus und unterziehen die kanonische Störungstheorie skalarer Felder einer Kritik um zu argumentieren, dass die kanonisch renormierte Quantenfeldtheorie Haags Theorem umgeht, da sie genau diese Bedingung nicht erfüllt. Weil kanonische Quantenfelder mathematisch nicht wohldefiniert sind, lässt sich dies nicht beweisen. Wir begnügen uns daher mit einem heuristischen Argument, das wir nichtsdestotrotz für überzeugend halten.

Wir sind der Auffassung, dass nichtstörungstheoretische Gleichungen für eine Definition von Quantenfeldern notwendig sind. Der Hopfalgebraische Zugang zur perturbativen Quantenfeldtheorie bietet die Möglichkeit, Dyson-Schwinger-und Renormierungsgruppengleichungen mathematisch sauber herzuleiten, wenn auch mit rein kombinatorischem Ausgangspunkt. Wir präsentieren eine Beschreibung dieser Methode und diskutieren eine gewöhnliche Differentialgleichung für die anomale Dimension des Photons. Eine Spielzeugmodellversion dieser Gleichung lässt sich exakt lösen; ihre Lösung weist eine interessante nichtstörunsgtheoretische Eigenschaft auf, deren Auswirkungen auf die laufende Kopplung und die Selbstenergie des Photons wir untersuchen. Solche nichtperturbativen Beiträge mögen die Existenz eines Landau-Pols ausschliessen, ein Sachverhalt, den wir ebenfalls diskutieren.

Unter der Arbeitshypothese, dass die anomale Dimension eines Quantenfeldes in die Klasse der resurgenten Funktionen fällt, studieren wir, welche Bedingungen die Dyson-Schwinger-und Renormierungsgruppengleichungen an ihre Transreihe stellen. Wir stellen fest, dass diese unter bestimmten Bedingungen kodieren, wie der perturbative Sektor den nichtperturbativen vollständig determiniert.

Contents

Introduction

Quantum field theory (QFT) is undoubtedly one of the most successful physical theories. Besides the often cited extraordinary precision with which the anomalous magnetic moment of the electron had been computed in quantum electrodynamics (QED), this framework enabled theorists to *predict the existence of hitherto unknown particles.*

As Dirac was trying to make sense of the negative energy solutions of the equation which was later named after him, he proposed the existence of a positively charged electron [**Schw94**]. This particle, nowadays known as the positron, presents an early example of a so-called *antiparticle.* It is fair to say that it was the formalism he was playing with that *led* him to think of such entities. And here we have theoretical physics at its best: the formulae under investigation only make sense provided an entity so-and-so exists. Here are Dirac's words:

> "A hole, if there were one, would be a new kind of particle, unknown to experimental physics, having the same mass and opposite charge to the electron. We may call such a particle an anti-electron. We should not expect to find any of them in nature, on account of their rapid rate of recombination with electrons, but if they could be produced experimentally in high vacuum, they would be quite stable and amenable to observations"[1].

Of course, the positron was not the only particle to be predicted by quantum field theory. W and Z bosons, ie the carrier particles of the weak force, both bottom and top quark and probably also the Higgs particle are all examples of matter particles whose existence was in some sense necessitated by theory prior to their discovery.

Yet canonical QFT presents itself as a stupendous and intricate jigsaw puzzle. While some massive chunks are for themselves coherent, we shall see that some connecting pieces are still only tenuously locked, though simply taken for granted by many practising physicists, both of phenomenological *and* of theoretical creed.

Constructive and axiomatic quantum field theory

In the light of this success, it seems ironic that so far physically realistic quantum field theories like the standard model (SM) and its subtheories quantum electrodynamics (QED) and quantum chromodynamics (QCD) all defy a mathematically rigorous description [**Su12**].

Take QED. While gauge transformations are classically well-understood as representations of a unitary group acting on sections of a principle bundle [**Blee81**], it is not entirely clear what becomes of them once the theory is quantised [**StroWi74, Stro13**]. However, Wightman and Gårding have shown that the quantisation of the free electromagnetic field due to Gupta and Bleuler is mathematically consistent in the context of Krein spaces (see [**Stro13**] and references there, p.156).

Drawing on the review article [**Su12**], we make the following observations as to what the state of affairs broadly speaking currently is.

[1]P.Dirac: *Quantised Singularities in the Electromagnetic Field.* Proc. R. Soc. Lond. A 133, 60-72 (1931)

- FIRST, all approaches to construct quantum field models in a way seen as mathematically sound and rigorous employ methods from operator theory and stochastic analysis, the latter only in the Euclidean case.

This is certainly natural given the corresponding heuristically very successful notions used in Lagrangian quantum field theory and the formalism of functional integrals.

These endeavours are widely known under the label *constructive quantum field theory*, where a common objective of those approaches was to obtain a theory of quantum fields with some reasonable properties. *Axiomatic quantum field theory* refined these properties further to a system of axioms. Several more or less equivalent such axiomatic systems have been proposed, the most prominent of which are:

(1) the so-called Wightman axioms [**StreatWi00, Streat75**],
(2) their Euclidean counterparts Osterwalder-Schrader axioms [**OSchra73, Stro13**] and
(3) a system of axioms due to Araki, Haag and Kastler [**HaKa64, Ha96, Stro13**].

These axioms were enunciated in an attempt to clarify and discern what a quantum field theory should or could reasonably be.

In contrast to this, the proponents of the somewhat idiosyncratic school of *axiomatic S-matrix theory* tried to discard the notion of quantum fields all together by setting axioms for the S-matrix [**Sta62**]. However, it lost traction when it was trumped by QCD in describing the strong interaction and later merged into the toolshed of string theory [**Ri14**].

- SECOND, efforts were made in two directions. In the constructive approach, models were built and then proven to conform with these axioms [**GliJaf68, GliJaf70**], whereas on the axiomatic side, the general properties of quantum fields defined in such a way were investigated under the proviso that they exist.

Among the achievements of the axiomatic community are rigorous proofs of the PCT and also the spin-statistics theorem [**StreatWi00**].

- THIRD, within the constructive framework, the first attempts started with superrenormalisable QFTs to stay clear of ultraviolet (UV) divergences.

The emerging problems with these models had been resolved immediately: the infinite volume divergences encountered there were cured by a finite number of subtractions, once the appropriate counterterms had been identified [**GliJaf68, GliJaf70**].

- FOURTH, however, these problems exacerbated to serious and to this day unsurmountable obstructions as soon as the realm of renormalisable theories was entered.

In the case of $(\phi^4)_d$, the critical dimension turned out to be $d = 4$, that is, rigorous results were attained only for the cases $d = 2, 3$. The issue there is that UV divergences cannot be defeated by a finite number of substractions. To our mind, it is their 'prolific' nature which lets these divergences preclude any nonperturbative treatment in the spirit of the constructive approach. For this introduction, suffice it to assert that a nonperturbative definition of renormalisation for renormalisable fields is clearly beyond constructive methods of the above type.

In particular, the fact that the regularised renormalisation Z factors can only be expected to have asymptotic perturbation series is obviously not conducive to their rigorous treatment. Although formally appearing in nonperturbative treatments as factors, they can a priori only be defined in terms of their perturbation series.

However, for completeness, we mention [**Schra76**] in which a possible path towards the (in some sense implicit) construction of $(\phi^4)_4$ in the context of the lattice approach was discussed. As one might expect, the remaining problem was to prove the existence of the renormalised limit to the continuum theory.

Haag's theorem and other triviality results

Around the beginning of the 1950s, soon after QED had been successfully laid out and heuristically shown to be renormalisable by its founding fathers [**Dys49b**], there was a small group of mathematical physicists who detected inconsistencies in its formulation.

Their prime concern turned out to be the *interaction picture* of a quantum field theory [**vHo52, Ha55**]. In particular, Haag concluded that it cannot exist unless it is trivial, ie only describing a free theory. Rigorous proofs for these suspicions could at the time not be given for a simple reason: in order to prove that a mathematical object does not exist or that it can only have certain characteristics, one has to say and clarifiy what kind of mathematical thing *it* actually is or what it is supposed to be.

But the situation changed when QFT was put on an axiomatic footing by Wightman and collaborators who made a number of reasonable assumptions and proved that the arguments put forward earlier against the interaction picture and *Dyson's matrix* were well-founded [**WiHa57**]. This result was then called *Haag('s) theorem*. It entails in particular that if a quantum field purports to be unitarily equivalent to a free field, it must be free itself.

Other important issues were the canonical (anti)commuation relations and the ill-definedness of quantum fields at sharp spacetime points. The ensueing decade brought to light a number of triviality results of the form "If X is a QFT with properties so-and-so, then it is trivial", where 'trivial' comes in 3 types, with increasing strength: the quantum fields are free fields, identity operators or vanishing. We shall see examples of all three types in this work.

The alternative formalism involving path integrals, although plagued by ill-definedness from the start [**AlHoMa08**], proved to be viable for lower spacetime dimensions in a Euclidean formulation [**GliJaf81**]. Schrader showed that a variant of Haag's triviality verdict also emerges there, albeit in a somewhat less devastating form [**Schra74**].

What to make of it

But in the light of the above-mentioned success of quantum field theories, the question about what to make of it is unavoidable. People found different answers.

On the physics side, the no-go results were widely ignored (apart from a few exceptions) or misunderstood and belittled as mathematical footnotes to the success story that QFT surely is (there will be quotes in the main text). Only confirming this, the author had several conversations with practising theoretical physicists (young and middle-aged) who had never heard of Haag's theorem and pertinent results[2].

On the mathematical physics side, the verdict was accepted and put down to the impossibility to implement relativistic quantum interactions in Fock space. And indeed, without much mathematical expertise, the evidence is clear: the UV divergences encountered in perturbation theory leave no doubt that something must be utterly wrong. Some were of the opinion that renormalisation only distracted the minds away from trying to find an appropriate new theory, as Buchholz and Haag paraphrase Heisenberg's view in [**BuHa00**].

Our philosophical stance on this is that *renormalised quantum field theory*, despite being a puzzle, provides us with *peepholes* through which we are allowed to glimpse at least some parts of that 'true' theory. Moreover, renormalisation follows rules which have a neat underlying algebraic structure and are not those of a random whack-a-mole game.

The title of this thesis started out as a working title. As the author's thoughts on this issue evolved but the answers he was trying to find were still vague and unclear, he obtained a few nonperturbative results so that suddenly, the title began to seem a bit narrow.

[2]We do not claim this to be representative, but we believe it is.

Yet it was decided that it is an apt title, if one lets 'Haag's theorem' stand for the triviality results that preclude a mathematically rigorous nonperturbative definition of interacting quantum field theories.

There are a vast number of more or less viable attempts to give QFT a sound mathematical meaning. Neither did we have the space nor the expertise to do all of them justice and include them here in our treatment. We have therefore chosen to direct our focus on the axiomatic approach: Haag's theorem was not just first formulated in this context [**StreatWi00**] but it is, as we find, conceptually closer to the canonical Lagrangian theory than any other competing formalism.

It goes without saying that beyond perturbation theory, *nonperturbative descriptions* have to be an integral part of any endeavour to characterise a QFT. We believe that studying nonperturbative equations like Dyson-Schwinger and renormalisation group equations should be part of the ongoing quest for understanding QFT. Our contributions in this direction worked out here are to be understood in this wider context: if we can one day prove that these nonperturbative equations have physical solutions whose properties adhere to a set of suitable axioms, then the reconstructed field theory is what the quest was about.

Outline

Each chapter begins with a detailed description of its content. We will in the following give an overview of the issues covered in this work and thereby explain our contribution.

Chapter 1 takes the reader on a journey through the history of Haag's theorem and some ensueing developments pertaining mainly to scalar theories. Experts in axiomatic quantum field theory will find a compendium of the bits and pieces they already know, while all other readers will learn of some interesting aspects from axiomatic quantum field theory.

The material we have garnered here includes several versions of Haag's theorem and related triviality results found by Wightman, Baumann, Powers, Strocchi and others. They provide enough evidence that a mathematically reasonable implementation of interactions in Fock space, ie the Hilbert space of a free field, is impossible. Along the way, we review and critique the arguments those authors used.

We discuss in particular the (anti)commutation relations, which originated in the Heisenberg uncertainty principle and whose role we find unclear in QFT. Although these relations are constitutive for free fields and the concept of particles in Fock space, they lose their meaning in interacting quantum field theories.

The link between renormalised perturbation theory and the scattering theory of Lehmann, Symanzik and Zimmermann is in our view still tenuous. We argue that the crux lies in what is known as the *wave-function renormalisation constant* Z, an object of questionable nonperturbative status, to put it mildly. We clarify that it cannot possibly satisfy the absurd condition of taking values in the unit interval $[0, 1] \subset \mathbb{R}$ and should for the time being better be constrained to its role as formal power series in the renormalisation of perturbative QFT.

Chapter 2 presents and scrutinises the axioms of Wightman and Gårding and discusses the proof of Haag's theorem and its provisos at length. A key element is the smearing in space and time with respect to test functions. We motivate this by discussing a very insightful triviality theorem due to Wightman where one can nicely see how quantum fields are doomed to a trivial existence if overfraught with conditions: while free fields exist at sharp times, this seems to be too strong a requirement for interacting fields.

According to Strocchi's results on gauge theories, the Wightman framework seems inapt for gauge fields. We review his results and describe the severe problems the axiomatic approach encounters here. Because Haag's theorem relies on the Wightman axioms, it does not apply directly to gauge theories. Nonetheless, the situation is no better for QED and QCD.

Pervading our exposition is the belief that the results obtained from renormalised QFT give us at least some vague hints about the features the sought-after theory should have. Because the *spectral condition* does not seem to be satisfied, especially in QED, where *spacelike* photons are a key concept, we suggest that this part of Wightman's axioms is questionable for QED.

Chapter 3 reviews the canonical derivation of the interaction picture and the Gell-Mann-Low formula which is the key identity attacked head-on by Haag's theorem. We then go through the folklore of renormalisation, which, in its canonical form, resembles more a narrative than a theory. The most important lesson that Haag's theorem teaches physics is in our mind that the renormalised theory cannot be unitarily equivalent to a free theory. Funnily enough, the general impression the author got from studying the literature was that physicists would very much like to retain precisely this property.

We show that the canonical procedure almost surely destroys this very feature. One flank of the argument is provided by a little theorem on free fields which can be found in [**ReSi75**]. It contains a simple truth: two free scalar fields of different masses are not unitarily equivalent. We have therefore dubbed this assertion 'Haag's theorem for free fields'. The proof uses none of the mathematically elaborated arguments involved in the proof of Haag's theorem, it is so simple and nontechnical that no physicist dare dismiss it as a purely mathematical fancy! The second flank is a nice standard canonical computation in which a massive free field is perturbed by a mass-shift interaction term in its Lagrangian. The resulting field is then also a free field, albeit with a different mass.

Although the map that takes the unperturbed free field to the perturbed one is within the canonical framework clearly portrayed as a unitary intertwiner, it will then be clear that this cannot be the case by Haag's theorem for free fields. Because the Lagrangian of a renormalised QFT also encompasses mass-shift interaction terms, the heuristic evidence leaves no room for any other conclusion than that the quantum field of a renormalised theory is unitarily inequivalent to a free field. This entails that however canonical perturbation theory might be interpreted mathematically, the central provision of unitary equivalence employed in Haag's theorem is almost surely *violated*. Therefore, Haag's theorem is not applicable to a renormalised quantum field theory and consequently, *renormalisation circumvents Haag's theorem*.

Chapters 4 & 5 serve as an introduction the combinatorial Hopf-algebraic approach to perturbative QFT and the two most important nonperturbative equations, namely Dyson-Schwinger and the renormalisation group equations. These chapters are written and included in this work for purely pedagogical reasons. Experts will find nothing new. Nevertheless, the material presented here is of value to those not in the know. As far as we can tell, the concepts discussed here are scattered over several research papers which often explore many mathematically interesting interconnections. Here we focus strictly on those aspects absolutely necessary to understand the nonperturbative results presented in the ensuing two chapters.

However, apart from preparing the ground for the last two chapters, our exposition contains some small aspects we have not found discussed in those papers, partly because they were obvious to the authors[3].

Chapter 6 investigates an ordinary differential equation for the anomalous dimension of the photon and studies a toy model approximation with an interesting nonperturbative feature. This material, which builds on earlier work by Kreimer and Yeats, had been previously published in [**KlaKrei13**] but is presented here in a more pedagogical form making it much more convenient to read. A small sign error in the spectral representation of the photon's self-energy in the original publication has been spotted and corrected.

Chapter 7 addresses some relatively novel themes concerning resurgent transseries and their possible applications in QFT. We make the working assumption that the anomalous dimension

[3]Examples are Dyson-Schwinger equations in the quotient Hopf algebra of QED and the Hopf-algebraic renormalisation of overall convergent Feynman diagrams.

is a resurgent function of the coupling and that it can therefore be represented by a resurgent transseries. We scrutinise what condition Dyson-Schwinger and renormalisation group equations introduced in Chapter 4 and 5 impose on the anomalous dimension's transseries and in particular how its perturbative sector is linked to its nonperturbative sectors. It turns out that on certain assumptions, the perturbative sector determines the nonperturbative one completely.

To our knowledge, this is the first time that transseries have been employed to study Dyson-Schwinger and renormalisation group equations. We have become aware of some very recent developments in string theory where transseries have been used to study differential equations only *after* our work on transseries had already been finished. We asked ourselves the question whether the little algebraic apparatus we had developed to tackle the problem was a bit over the top on account of its mathematical formality.

But as far as we can tell, we think that the complexity of the nonperturbative equations in question here fully vindicate our tools. The point is, simply inserting a transseries into these equations merely produces an indecipherable clutter, hard to handle and extract information from. The mathematically neat treatment, on the other hand, allows us to do just that.

The **appendix** contains in Part A a collection of preliminary mathematical background material that readers more or less need in order to follow the arguments in the main text. Part B has everything that we thought too technical for the main text which is why we have relegated it there.

CHAPTER 1

The representation issue and Haag's theorem

We take a historical journey and describe the developments that led up to Haag's theorem. Contrary to what its name suggests, it is the result of a collective effort and not of a single author. Yet it was Haag who put out the seminal paper in which some of the mathematical problems of the canonical formalism were first circumscribed, in particular those associated to the interaction picture representation of quantum field theory (QFT).

Because QFT was developed from nonrelativistic quantum mechanics and its conceptional foundations, it has inherited a bunch of ideas from it. We mention the canonical (anti) commutation relations (CCR/CAR) for observables and the time evolution of states by a one-parameter family of unitary operators. Concerning the CCR, it was already known that the Stone-von Neumann theorem cannot be applied to systems with an infinite number of degrees of freedom. This sparked doubts about whether the procedure of canonical quantisation picked the appropriate representation of the CCR and whether the interaction picture representation can actually be unitarily equivalent to the Heisenberg picture representation.

Section 1.1 depicts the corresponding representation issue and how it was handled at the time to introduce the backdrop for Haag's seminal publication [**Ha55**] whose salient points are covered in the first part of Section 1.2. The second part of Section 1.2 is devoted to the subsequent developments that culminated in Haag's theorem. Its proof, being rather technical, is deferred to Chapter 2.

The ensueing decades witnessed a number of pertinent results which we describe in Section 1.3. Of particular importance is an analogue of Haag's theorem in the Euclidean realm proven for a class of superrenormalisable theories, which we survey in Section 1.4. We will see there how the triviality dictum of Haag's theorem coexists peacefully with its very circumvention by (super)renormalisation!

Section 1.5 reviews the Fock space for an infinite number of degrees of freedom and has a critique of the interaction picture. Altogether, the arguments presented there provide compelling reasons why the interaction picture must be a fallacious business.

Section 1.6 presents the no-interaction theorems of Powers and Baumann. The central outcome there is that field theories conforming with the CCR/CAR must be necessarily free if the dimension of spacetime exceeds a certain threshold ($d \geq 3$ for fermions and $d \geq 5$ or $d \geq 4$ for bosons). Although free fields clearly satisfy these relations, this calls into question their meaning in a general QFT.

To make the case against the CCR/CAR, some authors bring in the field-strength (or wave-function) renormalsation constant. Since we deem this issue worthy of discussion, we have included some observations about this truely dubitable object in Section 1.7. We argue that it is not at all understood and only obstructs insight into the connection between asymptotic scattering theory and renormalised perturbation theory.

Finally, we close this chapter with Section 1.8, where we survey the reactions that Haag's theorem stirred among a minority of the physics community.

1.1. Inequivalent representations

Right at the outset, when quantum mechanics came into being in the 1920s, there was what one may call the *representation problem.* At the time, no one saw that the two competing formalisms - wave mechanics as developed by Schrödinger and matrix mechanics put forward by Heisenberg, Born and Jordan - were in fact equivalent.

Yet their proponents hardly appreciated each other's work. In 1926, Einstein wrote in a letter to Schrödinger, that he was convinced "that you have made a decisive advance with your quantum condition, just as I am equally convinced that the Heisenberg-Born route is off the track". Soon after, Schrödinger remarked in a note to a paper that their route left him "discouraged, if not repelled, by what appeared ... a rather difficult method ... defying visualisation" [**Rue11**]. And Heisenberg told Pauli, "the more I think of the physical part of the Schrödinger theory, the more detestable I find it. What Schrödinger writes about visualisation makes scarcely any sense, in other words I think it is shit[1]."[**Rue11**]

Canonical commutation relations. However, both formalisms had something in common: they were dealing with an algebra generated by operators $\{Q_1, ..., Q_n\}$ and $\{P_1, ..., P_n\}$ corresponding to the canonical position and momentum variables of Hamiltonian mechanics, which satisfy the *canonical commutation rules* (or *relations*)

$$[Q_j, Q_l] = 0 = [P_j, P_l] \ , \qquad [Q_j, P_l] = i\delta_{jl} \qquad \text{(CCR)} \tag{1.1.1}$$

for all $j, l \in \{1, ..., n\}$ on a Hilbert space $\mathfrak{H}$. In Heisenberg's matrix mechanics, these objects are matrices with infinitely many entries, whereas in Schrödinger's wave mechanics they are represented by the two operators

$$(Q_j\psi)(x) = x_j\psi(x) \ , \qquad (P_l\psi)(x) = -i\partial_l\psi(x) \tag{1.1.2}$$

which act on square-integrable wavefunctions $\psi \in L^2(\mathbb{R}^n) = \mathfrak{H}$ (see any textbook on quantum mechanics, eg [**Strau13**]).

1.1.1. Stone-von Neumann theorem. The dispute over which theory was the right one was settled when von Neumann took the cues given to him by Stone and proved in 1931 that both formulations of quantum mechanics are *equivalent* in the sense that both are *unitarily equivalent representations of the canonical commutation rules* (1.1.1) if their exponentiations

$$U(a) = \exp(ia \cdot P) \ , \qquad V(b) = \exp(ib \cdot Q) \qquad a, b \in \mathbb{R}^n \tag{1.1.3}$$

are so-called *Weyl unitaries* [**vNeu31**]. In the case of the Schrödinger representation, these Weyl unitaries are given by the two families of bounded operators defined as

$$(U_S(a)\psi)(x) = \psi(x + a) \qquad (V_S(b)\psi)(x) = e^{ib\cdot x}\psi(x) \tag{1.1.4}$$

for $\psi \in L^2(\mathbb{R}^n)$. The CCR (1.1.1) now take what is called the *Weyl form of the CCR*,

$$U_S(a)V_S(b) = e^{ia\cdot b}V_S(b)U_S(a) \qquad \text{(Weyl CCR).} \tag{1.1.5}$$

The Stone-von Neumann theorem makes the assertion that all Weyl unitaries conforming with these relations are unitarily equivalent to a finite direct sum of Schrödinger representations:

THEOREM 1.1 (Stone-von Neumann). *Let $\{U(a) : a \in \mathbb{R}^n\}$ and $\{V(b) : b \in \mathbb{R}^n\}$ be irreducible Weyl unitaries on a separable Hilbert space $\mathfrak{H}$, ie two weakly continuous families of unitary operators such that $U(a)U(b) = U(a + b)$, $V(a)V(b) = V(a + b)$ and*

$$U(a)V(b) = e^{ia\cdot b}V(b)U(a) \qquad \textit{(Weyl CCR)} \tag{1.1.6}$$

[1] "Ich finde es Mist". (see [**Strau01**])

for all $a, b \in \mathbb{R}^n$. Then there is a Hilbert space isomorphism $W : \mathfrak{H} \to L^2(\mathbb{R}^n)$ such that

$$WU(a)W^{-1} = U_S(a) \qquad WV(a)W^{-1} = V_S(a), \tag{1.1.7}$$

where U_S and V_S are the Schrödinger representation Weyl unitaries. If the above Weyl unitaries in (1.1.6) are reducible, then each irreducible subrepresentation is unitarily equivalent to the Schrödinger representation.

Proof. See [**vNeu31**] or any book on the mathematics of quantum mechanics, for example [**Em09**]. □

The reason why this theorem had to be phrased in terms of the Weyl CCR (1.1.6) and not the CCR is, as von Neumann pointed out in [**vNeu31**], that the CCR (1.1.1) can certainly not hold on the whole Hilbert space $L^2(\mathbb{R}^n)$ since the operators in (1.1.2) are *unbounded*. This is easy to see: $\int d^n x\ |\psi(x)|^2 < \infty$ does *not* imply $\int d^n x\ |x_j \psi(x)|^2 < \infty$. Moreover, if (1.1.1) were valid everywhere in the Hilbert space, one could take the trace of both sides yielding a contradiction[2].

Since both the CCR algebra of the Schrödinger and the Heisenberg representation of quantum mechanics generate irreducible Weyl unitaries, the issue was indeed settled. However, some questions remained:

- are there representations of the CCR (1.1.1) which do *not* generate Weyl unitaries and are therefore not unitarily equivalent to the Schrödinger and hence also not to the Heisenberg representation?
- On what conditions do they give rise to a representation of the Weyl CCR?

Dixmier [**Dix58**] found one particular answer to this latter question.

Theorem 1.2 (Dixmier). *Let Q, P be two closed symmetric operators on a Hilbert space $\mathfrak{H}$ with common stable domain $\mathfrak{D}$, ie $P\mathfrak{D} \subset \mathfrak{D}$ and $Q\mathfrak{D} \subset \mathfrak{D}$. Assume the operator*

$$H = P^2 + Q^2 \tag{1.1.8}$$

is essentially self-adjoint on $\mathfrak{H}$. If Q and P satisfy the CCR algebra

$$[Q, P] = i\ , \qquad [Q, Q] = 0 = [P, P] \tag{1.1.9}$$

on $\mathfrak{D}$ then $\mathfrak{H}$ decomposes into a direct sum of subspaces on each of which their restrictions are unitarily equivalent to the Schrödinger representation.

That $H = P^2 + Q^2$ is essentially self-adjoint seems physically reasonable as this operator corresponds to the Hamiltonian of the harmonic oscillator, the much beloved workhorse of quantum mechanics. In case the assumptions of Dixmier's theorem are not given, a number of examples for representations which are unitarily inequivalent to the Schrödinger representation have been found [**Su01**].

Surely, most examples of inequivalent representations are physically pathological. However, the interesting question is whether there is an example of physical relevance. And, yes, there is. Reeh found one such example [**Ree88**]: a (nonrelativistic, quantum mechanical) electron in the exterior of an infinitely long cylinder with a magnetic flux running through it.

To arrive at the model, one has to let the cylinder become infinitely thin, that is, in Reeh's description, become the z-axis. In doing so, he clearly stayed within the range of acceptable habits of a theoretical physicist.

Because the system is translationally invariant along the z-axis, there are only two degrees of freedom, ie in the above setting of (1.1.1) we have $n = 2$. The canonical momentum operators are

$$p_x = -i\partial_x + eA_1(x, y), \qquad p_y = -i\partial_y + eA_2(x, y), \tag{1.1.10}$$

[2] I thank David Broadhurst for pointing this out to me.

where $A_1(x, y)$ and $A_2(x, y)$ are the components of the electromagnatic vector potential and e is the electron's charge.

This particular example, although it satisfies the CCR, is *not unitarily equivalent to the Schrödinger representation.* Regarding Theorem 1.2, Reeh closes his paper by pointing out that $p_x^2 + p_y^2$ is not essentially self-adjoint, in agreement with Dixmier's result. What we learn from this is that

- FIRSTLY, even when the system has a finite number of degrees of freedom, not all representations of the CCR (1.1.1) are unitarily equivalent to the Schrödinger representation and
- SECONDLY, this need not worry us. It rather suggests that unitary equivalence is too strong a notion for physical equivalence.

Reeh's example suggests that we abandon the view that every quantum-mechanical system should lie in the unitary equivalence class of the Schrödinger representation.

In fact, a much weaker yet still physically sensible notion of equivalence has been put forward by Haag and Kastler in [**HaKa64**]. The authors essentially propose a form of weak equality of operators, namely that two observables A and B are equivalent if their matrix elements cannot be distinguished by measurement, that is, for a subset $\mathfrak{D}$ of state vectors, which describe the set of all possible experimental setups, one has

$$|\langle\Psi|(A - B)\Phi\rangle| < \varepsilon \qquad \forall\Psi, \Phi \in \mathfrak{D}, \tag{1.1.11}$$

in which $\varepsilon > 0$ is below any conceivable lower measuring limit[3].

1.1.2. Fock space. It was in one of the early papers on quantum field theory (QFT) in 1929 by Heisenberg and Pauli that the notion of what is nowadays known as *Fock space* first emerged [**HeiPau29**]. A bit later, this concept was explored more completely by Fock [**Fo32**] and rephrased in rigorous mathematical form by Cook [**Co53**]. This setting seemed to be appropriate and make sense even for relativistic particles. Because the Schrödinger representation of nonrelativistic many-particle systems can also be phrased in these terms, the Fock space became 'the Schrödinger representation of QFT' (see eg [**Di11**]).

Since Reeh's counterexample of a perfectly physical but nevertheless non-Schrödinger representation of the CCR in quantum mechanics was discovered rather late (1988!) and was not known at the time, the representation issue continued to be given plenty of attention.

It became topical again in the 1950s when Friedrichs constructed what he called *myriotic representations* of the CCR (1.1.1) on Fock space, also known as 'strange representations'. These representations are defined by the absence of any number operator and are obtained by passing to the limit of a countably infinite number of degrees of freedom [**Fried53**], ie $n \to \infty$ in (1.1.1). The Stone-von Neumann theorem is in this case no longer applicable.

In 1954, Wightman and Gårding published results proving that for this limit, there exists, as they put it, a "maze of irreducible inequivalent representations" [**GaWi54**]. However, we shall see in Section 1.6 that neither the CCR for bosons nor the anticommutation relations (CAR) for fermions are features that fully interacting theories can be expected to possess, at least in $d \geq 4$ spacetime dimensions. Therefore, the representation problem may actually be a *pseudo problem.* Of course, in the case of quantum mechanics, the CCR express the fundamental *Heisenberg uncertainty principle* which one is not willing to abandon. However, since there is no analogue of the position operator in QFT, it is not clear how this principle can be implemented through the CCR or CAR in a relativistic quantum theory[4]!

[3] The mathematical notion behind this is that of a convex topology induced by a system of *seminorms*: each pair of elements in $\mathfrak{D}$ defines a seminorm.

[4] Although for example the energy-time uncertainty is generally presumed to be true and employed in the interpretation of virtual off-shell particles in Feynman diagrams, there is no obvious connection to the CCR/CAR in QFT.

Van Hove phenomenon. Prior to these developments, van Hove was one of the first authors who tried to rigorously define a Hamiltonian of a massive interacting scalar field ϕ. The interaction he studied consists of a finite number of fixed point sources [**vHo52**], the Hamiltonian being

$$H_g = H_0 + gH_I, \tag{1.1.12}$$

where $H_0 = \sum_{\mathbf{k}} \epsilon_{\mathbf{k}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}}$ is the free and $H_I = \sum_{s=1}^{l} \beta_s \phi(\mathbf{x}_s)$ the interacting part. The point sources sit at positions $\mathbf{x}_s$ with strength $g\beta_s$. Introducing a momentum cutoff $\kappa > 0$ such that $\epsilon_{\mathbf{k}} = 0$ if $|\mathbf{k}| > \kappa$, he considered the two vacuum states Φ_0 and $\Phi_0(g)$ of H_0 and H_g, respectively, and found for their overlap

$$\langle \Phi_0 | \Phi_g \rangle \to 0 \qquad \text{as} \quad \kappa \to \infty \qquad\qquad \text{('van Hove phenomenon')}, \tag{1.1.13}$$

ie when the cutoff was removed by taking the limit, the Hilbert spaces of states turned out to be orthogonal for H_0 and H_g. He also found this to be the case for the energy eigenstates

$$H_g \Phi_n(g) = E_n(g) \Phi_n(g) \tag{1.1.14}$$

of energy $E_n(g)$ for different values of g, ie $\langle \Phi_n(g') | \Phi_m(g) \rangle \to 0$ as the cutoff was removed, for all n, m and $g' \neq g$.

He concluded that "the stationary states of the field interacting with the sources are no linear combinations of the stationary states of the free theory". A first sign that something may be wrong with the interaction picture, as Coleman wrote in a short review of van Hove's paper "it suggests that there is no mathematical justification for using the interaction representation and that the occasional successes of renormalization methods are lucky flukes ..."[5] .

1.2. Haag's theorem and its history

Such was the backdrop against which Haag argued in his seminal publication [**Ha55**] that the interaction picture cannot exist. The salient points he made were the following.

FIRST, it is very easy to find strange representations of the CCR (1.1.1) in the case of infinitely many degrees of freedom: a seemingly innocuous transformation like

$$q_\alpha \mapsto \widetilde{q}_\alpha = c q_\alpha, \qquad\qquad p_\alpha \mapsto \widetilde{p}_\alpha = c^{-1} p_\alpha \tag{1.2.1}$$

for any $c \notin \{0, 1\}$ of the canonical variables $\{q_\alpha, p_\alpha\}$ leads to a strange representation of the CCR, ie a representation for which there is no number operator and no vacuum state.

SECOND, Dyson's matrix $V = \mathsf{U}(0, -\infty)$ *cannot* exist, ie the operator that evolves interaction picture states from the infinitely far past at $t = -\infty$, where the particles are free, to the present at $t = 0$, where they (may) interact.

1.2.1. Strange representations. As regards the first point, let us see why (1.2.1) produces a strange representation. We follow Haag and write $c = \exp(\varepsilon)$ with $\varepsilon \neq 0$. For the annihilators and creators,

$$a_\alpha = \frac{1}{\sqrt{2}}(q_\alpha + ip_\alpha), \qquad\qquad a^\dagger_\alpha = \frac{1}{\sqrt{2}}(q_\alpha - ip_\alpha) \tag{1.2.2}$$

this transformation takes the form of a 'Bogoliubov transformation':

$$\begin{aligned} a_\alpha &\mapsto \widetilde{a}_\alpha = \cosh\varepsilon \ a_\alpha + \sinh\varepsilon \ a^\dagger_\alpha \\ a^\dagger_\alpha &\mapsto \widetilde{a}^\dagger_\alpha = \sinh\varepsilon \ a_\alpha + \cosh\varepsilon \ a^\dagger_\alpha, \end{aligned} \tag{1.2.3}$$

which is easy to check. We let $\alpha = 1, 2, ..., N$, where $N < \infty$ for the moment and $\mathfrak{H}$ be the Fock space generated by applying the elements of the CCR algebra $\langle a_\alpha, a^\dagger_\alpha : \alpha \in \{1, ..., N\} \rangle_{\mathbb{C}}$ to the

[5] Coleman's review of van Hove's paper is available at www.ams.org/mathscinet, keywords: author "van Hove", year 1952.

vacuum state. 'Strangeness' of the representation $\{\widetilde{a}_\alpha, \widetilde{a}_\alpha^\dagger\}$ will be incurred only when the limit $N \to \infty$ is taken. Notice that the generator of the above transformation (1.2.1) is given by

$$T = \frac{i}{2} \sum_{\alpha=1}^{N} [a_\alpha^\dagger a_\alpha^\dagger - a_\alpha a_\alpha], \tag{1.2.4}$$

ie $V = \exp(i\varepsilon T)$ is the transformation such that $\widetilde{a}_\alpha = V a_\alpha V^{-1}$ and $\widetilde{a}_\alpha^\dagger = V a_\alpha^\dagger V^{-1}$ which is unitary as long as N is finite. When $N \to \infty$, this operator will map any vector of the Fock representation to a vector with infinite norm. Let Ψ_0 denote the Fock vacuum of the original representation. Due to

$$\widetilde{a}_\alpha V \Psi_0 = V a_\alpha V^{-1} V \Psi_0 = V a_\alpha \Psi_0 = 0 \tag{1.2.5}$$

we see that the new representation does also have a vacuum. We denote it by $\widetilde{\Psi}_0 := V\Psi_0$. We abbreviate $\tau_\alpha := a_\alpha^\dagger a_\alpha^\dagger - a_\alpha a_\alpha$ and compute the vacua's overlap,

$$\langle \Psi_0 | \widetilde{\Psi}_0 \rangle = \langle \Psi_0 | V \Psi_0 \rangle = \langle \Psi_0 | (\prod_{\alpha=1}^{N} e^{-\frac{\varepsilon}{2}\tau_\alpha}) \Psi_0 \rangle = \prod_{\alpha=1}^{N} \langle \Psi_{\alpha,0} | e^{-\frac{\varepsilon}{2}\tau_\alpha} \Psi_{\alpha,0} \rangle, \tag{1.2.6}$$

where we have used that the vacuum is a tensor product $\Psi_0 = \Psi_{1,0} \otimes ... \otimes \Psi_{N,0}$ and $[\tau_\alpha, \tau_\beta] = 0$. Note that each factor in the product (1.2.6) yields the same value. This value is below 1 and it therefore vanishes in the limit $N \to \infty$, ie we find that the van Hove phenomenon occurs. Similiarly, one can show that, as Haag argues in [**Ha55**]

$$\langle \Psi | V \Phi \rangle = 0 \qquad \text{for all} \quad \Psi, \Phi \in \mathfrak{H} \qquad \text{(Fock space)} \tag{1.2.7}$$

in the limit. Although the new CCR algebra (1.2.3) is perfectly well-defined on $\mathfrak{H}$, its vaccum - if it exists - lies outside $\mathfrak{H}$! Hence *inside* $\mathfrak{H}$, this algebra is a 'strange representation', ie unitarily inequivalent to the Fock representation and has no vaccum.

The lesson Haag took from this was that seemingly minor and prima facie innocuous changes can easily lead to a theory which is unitarily equivalent when $N < \infty$, but ceases to be so in the limit $N \to \infty$.

Especially interesting is the vanishing overlap of the two vacua. If Dyson's matrix is well-defined for a finite number of degrees of freedom, then - as the above example shows - its existence is highly questionable in the case of an infinite system. A vanishing overlap of the two vacua directly contradicts what became known as the *theorem of Gell-Mann and Low* ([**GeMLo51**], cf. Section 3.1) which explicitly relies on a nonvanishing overlap: in the context of our example, the analogous statement is that up to a normalisation constant

$$\widetilde{\Psi}_0 = \frac{V\Psi_0}{\langle \Psi_0 | V \Psi_0 \rangle} \tag{1.2.8}$$

exists in the limit $N \to \infty$. Of course, the Bogoliubov transformation $V = e^{i\varepsilon T}$ as constructed in (1.2.4) bears no resemblance to Dyson's matrix. While Haag's example is therefore of no direct consequence for field theory, van Hove's indeed is: in his model,

$$\langle \Phi_n(g') | \Phi_m(g) \rangle = 0 \qquad (g \neq g') \tag{1.2.9}$$

implies that V vanishes weakly in $\mathfrak{H}$ since $\Phi_n(g') = V\Phi_n(g)$.

1.2.2. Dyson's matrix. However, Haag then went on to make the case that Dyson's matrix does not exist as follows ([**Ha55**], §4): let two Hermitian scalar fields $\phi_1(\mathbf{x})$ and $\phi_2(\mathbf{x})$ be related by a unitary map V according to

$$\phi_1(\mathbf{x}) = V^{-1} \phi_2(\mathbf{x}) V \tag{1.2.10}$$

and suppose their time evolution is governed by different Hamiltonians $H_1 \neq H_2$. Let $D(\mathbf{a})$ be a representation of the translation group in $\mathbb{R}^3$ under which both fields behave covariantly, ie

$$D(\mathbf{a})\phi_j(\mathbf{x})D(\mathbf{a})^\dagger = \phi_j(\mathbf{x}+\mathbf{a}) \qquad (j=1,2). \tag{1.2.11}$$

Then these conditions imply $[D^\dagger V^{-1}DV, \varphi_1] = 0$ from which $D^\dagger V^{-1}DV = 1$ and then also $[V, D(\mathbf{a})] = 0$ follow (by irreducibility of the field algebra[6]). This means $[V, \mathbf{P}] = 0$, where $\mathbf{P}$ is the three-component generator of translations in $\mathbb{R}^3$. Let Φ_{0j} be the vacuum of the representation of $\phi_j(\mathbf{x})$ $(j = 1, 2)$, ie $\Phi_{02} = V\Phi_{01}$. Then follows that V is trivial because

$$\mathbf{P}\Phi_{02} = \mathbf{P}V\Phi_{01} = V\mathbf{P}\Phi_{01} = 0 \tag{1.2.12}$$

implies $\Phi_{01} = w\Phi_{02}$ with $w \in \mathbb{C}$ since the vacuum is the only translation-invariant state. Haag now argues that this result is a contradiction to the assumption of both Hamiltonians having a different form, ie he means that $H_1\Phi_{01} = 0 = H_2\Phi_{01}$ is not acceptable if both Hamiltonians are to be different. In this view, the conclusion is that $H_1 = H_2$ and hence both theories are the same.

Mathematically, of course, this conclusion is not permissible if the two operators agree just on the vacuum. However, behind his statement "In all theories considered so far [the statement $\Phi_{01} = w\Phi_{02}$, author's note] is contradicted immediately by the form of the Hamiltonian."(ibidem) he refers to the more or less tacit assumption that one of the two operators should *polarise the vacuum*. This idea in turn originates in the fact that no one has ever seen an interacting Hamiltonian not constructed out of free fields, ie of annihilators and creators. All of those beheld by humans did always have a term of creators only, a term incapable of annihilating the vacuum.

Yet the above argument given by Haag against Dyson's matrix is flawed. If the two theories have different Hamiltonians, then they should also have different total momenta, ie in the language of Lagrangian field theory

$$\mathbf{P}_j = -\int d^3x \ \pi_j(\mathbf{x})\nabla\phi_j(\mathbf{x}), \tag{1.2.13}$$

and we expect $\mathbf{P}_1 = V^{-1}\mathbf{P}_2V$ (considering the complexity of the Poincaré algebra, the statement $\mathbf{P}_1 = \mathbf{P}_2$ is a rather strong assumption that needs discussion!). This suggest that the two fields should be covariant with respect to different representations of the translation group which means their generators $\mathbf{P}_1, \mathbf{P}_2$ are not the same and consequently (1.2.11) is bogus.

1.2.3. Results by Hall and Wightman. This is probably (we do not know) what Hall and Wightman had in mind when they wrote "In the opinion of the present authors, Haag's proof is, at least in part, inconclusive."(see [**WiHa57**], ref.10). But Haag did have a point there.

Hall and Wightman 'polished and generalised' Haag's argument (as they add) to obtain a result which at first glance seems less harmful to Dyson's matrix [**WiHa57**]: because both theories are different, one should allow for two Hilbert spaces $\mathfrak{H}_1$ and $\mathfrak{H}_2$, each equipped with a representation of the Euclidean group (rotations and translations) $D_j(\mathbf{a}, R)$ $(j = 1, 2)$ in $\mathbb{R}^3$ such that

$$D_j(\mathbf{a}, R)\phi_j(\mathbf{x})D_j(\mathbf{a}, R)^\dagger = \phi_j(R\mathbf{x}+\mathbf{a}) \qquad (j=1,2) \tag{1.2.14}$$

and assume there exist invariant states $\Phi_{0j} \in \mathfrak{H}_j$, that is, $D_j\Phi_{0j} = \Phi_{0j}$. Then it follows by irreducibility of both field theories from (1.2.10) that[7]

$$D_1(\mathbf{a}, R) = V^{-1}D_2(\mathbf{a}, R)V \qquad \text{and} \qquad V\Phi_{01} = \Phi_{02}. \tag{1.2.15}$$

[6]We will explain this thoroughly in Section 2.3, cf.(2.3.26).

[7]We discuss the proof more thoroughly in Section 2.3.

This seems to be a less devastating result because the map $V\colon \mathfrak{H}_1 \to \mathfrak{H}_2$ need not be trivial. Yet, as they subsequently showed, this cannot come to the rescue of Dyson's matrix either. If we just consider what it means for the n-point functions,

$$\begin{aligned}\langle \Phi_{01}|\phi_1(\mathbf{x}_1)...\phi_1(\mathbf{x}_n)\Phi_{01}\rangle &= \langle \Phi_{01}|V^{-1}\phi_2(\mathbf{x}_1)V...V^{-1}\phi_2(\mathbf{x}_n)V\Phi_{01}\rangle \\ &= \langle V\Phi_{01}|\phi_2(\mathbf{x}_1)V...V^{-1}\phi_2(\mathbf{x}_n)V\Phi_{01}\rangle \\ &= \langle \Phi_{02}|\phi_2(\mathbf{x}_1)...\phi_2(\mathbf{x}_n)\Phi_{02}\rangle,\end{aligned} \tag{1.2.16}$$

we see that they agree. This entails for the Heisenberg fields $\phi_j(t,\mathbf{x})$ that their n-point functions coincide on the time slice $t=0$. To extend this to a larger subset, let now the condition of Euclidean covariance of the Schrödinger fields in (1.2.14) be strengthened to *relativistic covariance*, ie Poincaré covariance,

$$U_j(a,\Lambda)\phi_j(x)U_j(a,\Lambda)^\dagger = \phi_j(\Lambda x + a) \qquad (j=1,2), \tag{1.2.17}$$

where $x=(t,\mathbf{x}), a \in \mathbb{M}$ are Minkowski spacetime points and Λ a proper orthochronous Lorentz transformation, then for pairwise spacelike-distant $x_1,...,x_n \in \mathbb{M}$ one has

$$\langle \Phi_{01}|\phi_1(x_1)...\phi_1(x_n)\Phi_{01}\rangle = \langle \Phi_{02}|\phi_2(x_1)...\phi_2(x_n)\Phi_{02}\rangle \qquad \text{(spacelike distances)}. \tag{1.2.18}$$

Hall and Wightman proved that for $n \leq 4$, this equality can be extended (in the sense of distribution theory) to all spacetime points, where $x_j \neq x_l$ if $j \neq l$. This result is referred to as *generalised Haag's theorem*. The term 'generalised' has been used because none of the fields need be free for (1.2.18) to hold (we have not required any of the fields to be free so far). The reason the authors could not prove this for higher n-point functions is that there is no element within the Poincaré group capable of jointly transforming $m \geq 4$ Minkowski spacetime vectors[8] to m arbitrary times but only to a subset which is not large enough for a complete characterisation [**StreatWi00**].

Notice that (1.2.16) holds for any quantum field, of whatever spin, it is a trivial consequence of (1.2.10) and the assumed irreducibility of the field algebra. The only difference for fields of higher spin is that the transformation law (1.2.17) needs a finite dimensional representation of the Lorentz group for spinor and vector fields and makes (1.2.18) less obvious. We shall discuss this point in Section 2.4 to see that it works out fine also in these cases.

The result (1.2.18) mattered and still matters because a field theory was by then already known to be sufficiently characterised by its vacuum expectation values, as shown by Wightman's *reconstruction theorem* put forward in [**Wi56**], which, in brief, says that a field theory can be (re)constructed from its vacuum expectation values. We shall survey this result in Section 2.2 and discuss the case of quantum electrodynamics (QED) in Section 2.4.

As the above arguments show, an interacting field theory cannot be unitarily equivalent to a free field theory unless we take the stance that it makes sense for an interacting field to possess n-point functions that for $n \leq 4$ agree with those of a free field.

1.2.4. Contributions by Greenberg and Jost. In fact, this stance has to be changed since a couple of years later, Greenberg proved in [**Gre59**] that if one of the two fields is a free field, then the equality (1.2.18) holds for all n-point functions and all spacetime points. His proof is inductive showing that if the n-point functions coincide for $n \leq 2m$, then they do so for $n \leq 2m+2$. It is to this day still an open question whether (1.2.18) is true for two general (Hermitian) fields at arbitrary spacetime points [**Streat75**].

Around the same time, using different arguments, another proof of Greenberg's result was obtained by Jost and Schroer [**Jo61**] and is therefore known under the label *Jost-Schroer theorem*: if a field theory has the same two-point function as a free field of mass $m > 0$, then that

[8] Recall that by translation invariance an n-point function is a function of $m = n-1$ Minkowski spacetime points.

is already sufficient for it to be a free field of the same mass. It was also independently shown by Federbush and Johnson [**FeJo60**] and the massless case was proved by Pohlmeyer [**Po69**].

1.2.5. Haag's theorem I. With these latter results, we arrive at what became known as Haag's theorem. Let us first state it in words, a more thorough exposition including the proof will be given in Section 2.3:

HAAG'S THEOREM. *If a scalar quantum field is unitarily equivalent to a free scalar quantum field, then, by virtue of the reconstruction theorem, it is also a free field because all vacuum expectation values coincide.*

As a consequence, Dyson's matrices, which purportedly transform in a unitary fashion the incoming and the outgoing free asymptotic fields into fully interacting fields, cannot exist. Note that the equality of the vacuum expectation values for spacelike separations (1.2.18) needs no other provisions than

- unitary equivalence of the two fields through the intertwiner V,
- Poincaré covariance and
- irreducibility of their operator algebras to warrant (1.2.15).

The remainder of the results, which say that the equality extends beyond spacelike separations into the entire Minkowski space $\mathbb{M}$, and their provisions that complete Haag's theorem, eg the Jost-Schroer theorem, therefore have a different status!

Because the proof of Haag's theorem is fairly technical and requires a number of assumptions, we defer it to Section 2.3. As the above deliberations suggest, the spacetime dimension does not enter the discussion anywhere and is therefore irrelevant.

Haag's theorem is a very deep and fundamental fact, true both for superrenormalisable and renormalisable theories. In its essence, it is rather trivial: it is a theorem about a free field of fixed mass and its unitary equivalence class. In fact, due to Theorem 3.1, which we call 'Haag's theorem for free fields', we know that even two free fields lie in distinct equivalence classes whenever their masses differ, however *infinitesimally* small that difference might be.

Galilean exemptions. Note that the step from equal-time vacuum expectation values (1.2.16) to (1.2.18) is not permitted for Galilean quantum field theories as employed in solid state physics. In fact, *Haag's theorem breaks down for Galilean quantum field theories* as the Jost-Schroer theorem does not hold for them: Dresden and Kahn showed that there are non-trivial (=interacting) Galilean quantum field theories whose two-point Wightman functions are identical to those of free fields [**DresKa62**]. We therefore have reason to believe that Haag's theorem in the strict sense of the above stated theorem is specific to relativistic quantum field theories.

However, *Euclidean quantum field theories* are a different case for which an analogue of Haag's theorem does indeed hold, as we shall see in Section 1.4 in the superrenormalisable case. How a nontrivial interacting theory is still attained despite this theorem's dictum, will also be shown there.

1.3. Other versions of Haag's theorem

In the ensuing decades, various related results were published. We shall survey them in this and the following sections. Before we discuss a very important variant of Haag's theorem for Euclidean field theories in the next section, we first have a look at two other versions of Haag's theorem. Especially the Streit-Emch theorem is worth being considered as it is closer to Haag's original formulation which we discussed in the previous section, Subsection 1.2.2.

1.3.1. Work by Emch & Streit. Emch presents a very different variant of Haag's theorem in his monograph [**Em09**] based on results proved by Streit in [**Strei68**].

Emch writes about the previously published proofs that "...these proofs, however, rely rather heavily on the analytic properties of the Wightman functions, which themselves reflect the locality and spectrum conditions, and tend to obscure the simple algebraic and group-theoretical facts actually responsible for the results obtained."([**Em09**], p.247).

Unlike the older versions, the Emch-Streit result focusses on the Weyl representations of the CCR generated by the field $\varphi(f)$ and its canonical momentum $\pi(g)$, smeared by a test functions f and g, ie

$$U(f) = e^{i\varphi(f)} \ , \qquad V(g) = e^{i\pi(g)} \tag{1.3.1}$$

which then satisfy the Weyl form of the CCR: $U(f)V(g) = e^{i(f,g)}V(g)U(f)$. This version of Haag's theorem is purportedly more general by assuming neither relativistic covariance nor causality (also known as locality, see Section 2.2).

What the authors assume instead is covariance with respect to a more general symmetry group that exhibits a property named 'η-clustering'. This feature of the symmetry, defined via some averaging process, is essentially the clustering property known for spacelike translations in relativistic theories (cf.(2.2.17) in Section 2.2). We have reason to believe, however, that for viable quantum field theories, the set of assumptions used in Emch's proof implies the very two conditions purportedly not needed, ie relativistic covariance and causality.

Before we elaborate on this point, let us have a look at the assertion of the Emch-Streit theorem. The upshot there is, interestingly, very close to that originally stated by Haag which we alluded to in Subsection 1.2.2, where we mentioned the polarisation of the vacuum: vacuum polarisation cannot occur if both Weyl representations of the CCR algebra are to be unitarily equivalent.

Let V be the unitary transformation connecting the fields $\varphi_1(f)$ and $\varphi_2(f)$ and their canonical conjugates, then the outcome is

$$H_2 = VH_1V^{-1} \tag{1.3.2}$$

for the corresponding generators of time translations, ie the Hamiltonians. Let H_1 be the Hamiltonian of the free field which exhibits no vacuum polarisation: $H_1\Psi_{01} = 0$. Then the other vacuum is also not polarised:

$$H_2\Psi_{02} = H_2V\Psi_{01} = VH_1\Psi_{01} = 0. \tag{1.3.3}$$

Because the decomposition $H_2 = H_1 + H_{int}$ with some interaction part H_{int} is not compatible with unitary equivalence (1.3.2) of both Hamiltonians, the authors conclude that the other theory is also free.

We shall not present the Streit-Emch theorem in its general form here as even its provisions are rather technical. The interested reader is referred to [**Em09**]. Instead, we quickly discuss a simpler version with a very elegant proof taken from [**Fred10**] which nevertheless shows that vacuum polarisation cannot occur in Fock space.

THEOREM 1.3 (No vacuum polarisation). *Let $\mathfrak{H}$ be the Fock space of a free field φ_0 with time translation generator H_0 and $U(\mathbf{a})$ a representation of the translation subgroup with invariant state $\Omega \in \mathfrak{H}$ (the vacuum). Assume there exists a field φ with well-defined sharp-time limits $\varphi(t,f)$ and $\partial_t\varphi(t,f) = \dot\varphi(t,f)$ for all test functions $f \in \mathscr{S}(\mathbb{R}^n)$ such that*

(i) $\varphi(0,f) = \varphi_0(0,f)$ *and* $\dot\varphi(0,f) = \dot\varphi_0(0,f)$;
(ii) $U(\mathbf{a})\varphi(t,f)U(\mathbf{a})^\dagger = \varphi(t,\tau_a f)$, *where* $(\tau_a f)(\mathbf{x}) = f(\mathbf{x}-\mathbf{a})$;
(iii) *there exists a self-adjoint operator H on $\mathfrak{H}$ with $[H,U(\mathbf{a})] = 0$ for all $\mathbf{a} \in \mathbb{R}^n$ and*

$$\varphi(t,f) = e^{iHt}\varphi(0,f)e^{-iHt}. \tag{1.3.4}$$

Then there is a constant $\lambda \in \mathbb{C}$ so that $H = H_0 + \lambda$ and thus $\varphi_0 = \varphi$.

PROOF. Because the free field generates a dense subspace in $\mathfrak{H}$, it suffices to show $H = H_0+\lambda$ on a state of the form $\Psi = \varphi_0(0, f_1)...\varphi_0(0, f_n)\Omega$ for test functions $f_1, ..., f_n$. The Heisenberg picture evolution (1.3.4) imlies $\dot{\varphi}(t, f) = i[H, \varphi(t, f)]$ and the condition $[H, U] = 0$ entails that the vacuum Ω is an eigenstate of H. Let λ be the corresponding eigenvalue, ie $(H - \lambda)\Omega = 0$. Then, by repeated application of $H\varphi(0, f) = \varphi(0, f)H - i\dot{\varphi}(0, f)$, we get

(1.3.5)
$$\begin{aligned}(H-\lambda)\varphi_0(0,f_1)...\varphi_0(0,f_n)\Omega &= (H-\lambda)\varphi(0,f_1)...\varphi(0,f_n)\Omega\\ &= -i\sum_{j=1}^{n}\varphi(0,f_1)...\dot{\varphi}(0,f_j)...\varphi(0,f_n)\Omega = -i\sum_{j=1}^{n}\varphi_0(0,f_1)...\dot{\varphi}_0(0,f_j)...\varphi_0(0,f_n)\Omega\\ &= H_0\varphi_0(0,f_1)...\varphi_0(0,f_n)\Omega.\end{aligned}$$

The assertion then follows straightforwardly from (1.3.4):

$$\varphi(t, f) = e^{iHt}\varphi(0, f)e^{-iHt} = e^{iH_0t}\varphi(0, f)e^{-iH_0t} = e^{iH_0t}\varphi_0(0, f)e^{-iH_0t} = \varphi_0(t, f). \tag{1.3.6}$$

□

We come back to the point made above concerning covariance and causality: contrary to what Emch and Streit claim, we believe that both relativistic covariance and causality are implied in their assumptions.

Lurking behind clustering, ie the well-known cluster decomposition property of vacuum expectation values, however, is causality, namely that the fields commute at spacelike distances[9]. The Emch-Streit theorem can only be more general for a model and not rely on relativistic covariance and causality, if there are symmetries other than translations with the same properties and only if, on top of that, lacking causality cannot obstruct the clustering feature. In other words, because the clustering property is a consequence of both causality and translational covariance, we deem it highly questionable that there is any symmetry in QFT other than translations that gives rise to clusters without the aid of causality.

Vacuum polarisation in Galilean theories. As presented by Lévy-Leblond in [**LeBlo67**], there are Galilean quantum field theories which provide nonrelativistic examples of interacting field theories that do *not* polarise the vacuum and thereby defy Haag's argument we presented in Subsection 1.2.2 regarding the polarisation of the vacuum. We briefly expound the author's argument.

Let $H, \{P_j, J_j, K_j\}$ be the generators of the Galilean group, ie the Galilean algebra consisting of the generators of time translations H, spatial translations P_j, rotations J_j and Galilean boosts K_j. In contrast to the Poincaré algebra, the subset $\{P_j, J_j, K_j\}$ is a Lie subalgebra which remains unaltered if the Hamiltonian is augmented by an interaction term. The crucial difference is displayed by the commutators

$$\text{Galilean case: } [K_j, P_l] = i\delta_{jl}m\ , \qquad \text{Poincaré case: } [K_j, P_l] = i\delta_{jl}H, \tag{1.3.7}$$

where $m > 0$ is the Galilean particle's mass and $\{K_j\}$ on the Poincaré side are the Lorentz boost generators in the Poincaré algebra. These commutators show clearly that altering the Poincaré algebra's Hamiltonian H requires the other generators be modified accordingly. More precisely, the point now is this: both Galilei and Poincaré algebra share the commutator

$$[K_j, H] = iP_j. \tag{1.3.8}$$

It is now possible in the Galilean case to find a perfectly physical interaction term H_I without messing up this commutator, ie (1.3.8) is left in peace because $[K_j, H_I] = 0$ [**LeBlo67**]. Even if one finds a Poincaré counterpart, it will interfere with the commutators in (1.3.7) which, as we said above, entails that the other generators must inevitably be changed as well. Let now

[9]This is discussed in Section 2.2, equation (2.2.17).

$H = H_0 + H_I$ be a Galilean Hamiltonian with free and interacting part H_0 and H_I, respectively, such that $[K_j, H_0] = [K_j, H] = iP_j$. Let Ψ_0 be the vacuum with $H_0\Psi_0 = 0$. Then

$$[K_j, H]\Psi_0 = iP_j\Psi_0 = 0 \qquad \Rightarrow \qquad K_j H\Psi_0 = HK_j\Psi_0 = 0, \tag{1.3.9}$$

because the vacuum is Galilean invariant. One infers from $K_j H\Psi_0 = 0$ that either the vacuum is polarised, $H\Psi_0 = \Psi_0$ (up to a prefactor) or not, ie $H\Psi_0 = 0$. If the former holds, then $H' := (H - \mathrm{id})$ will do the job and there is no vacuum polarisation.

Streit-Emch theorem more general? This result and the following thought calls the purported generality with respect to the symmetry group of the Emch-Streit theorem into question. One of the assumptions made by Emch and Streit is *cyclicity of the vacuum*[10]. As Fraser cogently explains in her thesis [**Fra06**], it is at least this property which Galilean field theories can easily violate. The standard example is a quantised Schrödinger field $\psi(t, \mathbf{x})$, defined as a solution of the second quantised Schrödinger equation

$$i\frac{\partial}{\partial t}\psi(t, \mathbf{x}) = -\frac{1}{2m}\Delta\psi(t, \mathbf{x}) \tag{1.3.10}$$

which has only positive-energy solutions of the form

$$\psi(t, \mathbf{x}) = \int d^3k \; e^{-i(\frac{\mathbf{k}^2}{2m}t - \mathbf{k}\cdot\mathbf{x})}c(\mathbf{k}) \tag{1.3.11}$$

and is not capable of creating a dense subspace, even when its adjoint is taken to join the game. Rather, they create subspaces known as *superselection sectors* for both particles with mass $m > 0$ and $-m < 0$ (see [**Fra06**] and references there[11]).

So, in summary, the cyclicity of the vacuum can probably only be achieved by relativistic fields. Granted that on the face of it, the Emch-Streit theorem is more general, the question as to what theories other than relativistic ones satisfy its assumptions if Galilean field theories fail to do so becomes even more pressing. Our conviction therefore is: none.

1.3.2. Algebraic version. Another version of Haag's theorem can be found in [**Wei11**] which we prefer to mention only briefly. The author proved this theorm in the context of *algebraic quantum field theory* as introduced by Haag and Kastler in [**HaKa64**] (see also [**Ha96, Bu00**] and [**Em09**]). For those readers acquainted with this somewhat idiosyncratic take on quantum field theory, here is very roughly speaking the upshot, unavoidably couched in algebraic language: the unitary equivalence class of a net of local von Neumann algebras is completely determined on a spacelike hyperplane. This means that if the von Neumann algebras of two field theories are related on a spacelike hypersurface through a unitary map, then they are unitarily equivalent and all vacuum expectation values agree.

1.3.3. Noncommutative QFT. The last version of Haag's theorem we mention here has been found in the context of *noncommutative quantum field theory* and is presented in [**AMY12**]. This peculiar theory purports to generalise QFT by letting the time coordinate fail to commute with the spatial coordinates. The result there says that if two theories are related by a unitary transformation and the S-matrix is trivial in one theory, so it must be in the other. The interested reader wanting to embark on a 'noncommutative journey' is referred to [**AMY12**] as a starting point.

[10] This will be defined in Section 2.2 as part of the Wightman axioms

[11] Fraser's text is an excellent piece of work devoted to Haag's theorem from a philosophical point of view.

1.4. Superrenormalisable theories evade Haag's theorem

We shall now discuss a result by Schrader which facilitates the understanding of the meaning of Haag's theorem for renormalisable QFTs profoundly. We know that Haag's theorem is independent of the dimension of spacetime and hence is also true for superrenormalisable QFTs. Because there is a well-known connection between Euclidean and relativistic quantum field theories [**OSchra73**], one has to expect there to be a Euclidean manifestation of Haag's theorem.

And indeed, this is exactly what Schrader's result says: an analogue of Haag's theorem holds also in the Euclidean realm. The good news is that superrenormalisable quantum theories are relatively well-understood courtesy of the work of constructive field theorists.

Although the protagonists did not necessarily think of it that way (they did not say), we shall now have a glimpse at their results and see that Haag's triviality dictum has in fact been proven to be evaded in a big class of superrenormalisable field theories by, one may say, *(super)renormalisation*!

1.4.1. Euclidean realm. Schrader proved in [**Schra74**] that Haag's theorem is also lurking in Euclidean field theories of the type $P(\varphi)_2$, where this symbol stands for the interaction term of the Hamiltonian in the form of a polynomial P bounded from below ('semi-bounded') with $P(0) = 0$ ('normalised').

In this context, a Euclidean field φ is a random variable with values in the set $\mathscr{D}'(\mathbb{R}^2)$ of distributions $\omega : \mathscr{D}(\mathbb{R}^2) \to \mathbb{C}$ on the space of test functions of compact support. Let us write them as

$$\omega(f) = \int d^2x \; \omega(x) f(x) \qquad\qquad f \in \mathscr{D}(\mathbb{R}^2). \tag{1.4.1}$$

The expectation values are given by *functional integrals*. For example, the expectation values of a free field are given by the functional integral

$$\langle \varphi(f)\varphi(h)\rangle_0 = \int_{\mathscr{D}'(\mathbb{R}^2)} \omega(f)\omega(h) \; d\mu_0(\omega), \tag{1.4.2}$$

where $f, h \in \mathscr{D}(\mathbb{R}^2)$ are test functions and μ_0 is the Gaussian measure with respect to the operator $-\Delta + m^2$. Then, given a semi-bounded nontrivial real and normalised polynomial P, the interacting part of the Euclidean action with coupling $\lambda > 0$ is defined by

$$V_{l,\lambda}(\varphi) := \lambda \int_{-l/2}^{+l/2} \int_{-l/2}^{+l/2} d^2x : P(\varphi(x)) :, \tag{1.4.3}$$

where : ... : stands for Wick ordering, the Euclidean analogue of normal ordering (see Subsection 3.3.1 or [**GliJaf81**]) and l is the box length which serves as a volume cutoff. Schrader builds on Newman's findings, namely that the family of measures given by

$$d\mu_{l,\lambda}(\omega) = \frac{e^{-V_{l,\lambda}(\omega)}}{\langle e^{-V_{l,\lambda}(\varphi)}\rangle_0} \; d\mu_0(\omega) \tag{1.4.4}$$

converges to a measure $\mu_{\infty,\lambda}$ for sufficiently small $\lambda > 0$ through the limit

$$\lim_{l'\to\infty} \int_{\mathscr{D}'(\mathbb{R}^2)} \omega(\chi_l f_1)...\omega(\chi_l f_n) \; d\mu_{l',\lambda}(\omega) = \int_{\mathscr{D}'(\mathbb{R}^2)} \omega(\chi_l f_1)...\omega(\chi_l f_n) \; d\mu_{\infty,\lambda}(\omega) \tag{1.4.5}$$

for all $l > 0$, where χ_l is a test function with compact support in the box. The f_j's are any test functions, ie the expectation values are well-defined for all test functions with compact support. Interestingly, Schrader's result is now that $\mu_{\infty,\lambda}$ and μ_0 have mutually disjoint support (for sufficiently small $\lambda > 0$). Thus, although the two-point function of the cutoff theory

$$\langle \varphi(f)\varphi(h)\rangle_{l,\lambda} = \int_{\mathscr{D}'(\mathbb{R}^2)} \omega(f)\omega(h)\ d\mu_{l,\lambda}(\omega) = \int_{\mathscr{D}'(\mathbb{R}^2)} \frac{e^{-V_{l,\lambda}(\omega)}}{\langle e^{-V_{l,\lambda}(\varphi)}\rangle_0}\ \omega(f)\omega(h)\ d\mu_0(\omega) \tag{1.4.6}$$

has a measure of the same support as the free measure, this situation changes dramatically in the limit to the full (ostensibly physical) theory when $l \to \infty$. While one may expect to be able to approximate (1.4.6) by employing perturbation theory with respect to λ to the last integral, the result of mutually disjoint measure support suggests that this method will lead to anything but an approximation. However, since the limit measure $\mu_{\infty,\lambda}$ exists for small enough $\lambda > 0$, the model was further explored with the following highly interesting results.

1. The Schwinger functions, ie the Euclidean Green's functions

$$S_\lambda(x_1, ..., x_n) := \langle \varphi(x_1)...\varphi(x_n)\rangle_\lambda = \int_{\mathscr{D}'(\mathbb{R}^2)} \omega(x_1)...\omega(x_n)\ d\mu_{\infty,\lambda}(\omega) \tag{1.4.7}$$

   exist in the sense of distributions. Let $f \in \mathscr{S}(\mathbb{R}^{2n})$ be a test function. Then the function $\lambda \mapsto S_\lambda(f)$ is smooth in λ in an interval $[0, \lambda_0)$ for $\lambda_0 > 0$ sufficiently small and has a Borel-summable asymptotic Taylor series [**Di73**].
2. Their Minkowski limit $\mathbb{R}^2 \ni (x^0, x) \to (it, x)$ yield distributions that satisfy the axioms due to Wightman [**GliJaSp74**] and give rise to a Wightman theory with nontrivial S-matrix [**OS76**], obtained through a nonperturbative LSZ reduction expansion[12].

What we see here is that although a Euclidean variant of Haag's theorem is valid, it does not preclude the existence of nontrivial interactions. Heuristically, it is easy to see that the measure $\mu_{\infty,\lambda}$ may actually be seen as a 'superrenormalised' measure: if we write

$$\langle e^{-V_{l,\lambda}(\varphi)}\rangle_0 = e^{-E_{l,\lambda}} \tag{1.4.8}$$

with ground state energy $E_{l,\lambda} = \langle V_{l,\lambda}(\varphi)\rangle_0$, then (1.4.4) becomes

$$d\mu_{l,\lambda}(\omega) = e^{-[V_{l,\lambda}(\omega) - E_{l,\lambda}]} d\mu_0(\omega). \tag{1.4.9}$$

This is exactly the kind of 'renormalisation' that Glimm and Jaffe used in their $(\varphi^4)_2$ model which they treated in the operator approach [**Jaff69, GliJaf68, GliJaf70**].

1.4.2. Evasion of Haag's theorem. If we combine Schrader's Euclidean result with the above points, we can draw a clear conclusion. Even though this result is about the class $P(\varphi)_2$, ie a big class of superrenormalisable quantum field models, it is of particular importance as it helps us understand what Haag's theorem may mean for renormalisable field theories.

The fact that the two measures μ_0 and $\mu_{\infty,\lambda} = \lim_{l\to\infty} \mu_{l,\lambda}$ have mutually disjoint support tells us that there exists no Radon-Nikodym density relating these two. This means that the Radon-Nikodym density in (1.4.9) ceases to make sense in the limit. In the operator approach, this probably corresponds to unitary inequivalence between the free and the interacting theory. Yet still: (super)renormalisation leads to a sensible result, ie another measure which describes a nontrivial theory.

This is what we opine about the meaning of Haag's theorem for renormalisable theories and try to make plausible in this work: once renormalised, these theories are nontrivial and unitary inequivalent to the very free theories employed to construct them. In other words, it is precisely renormalisation what allows us to stay clear of Haag's theorem.

[12] No concrete result for comparison with any perturbative result was computed. This is because one has to know the n-point functions to make practical use of the LSZ formula!

1.5. The interaction picture in Fock space

The following results provide very compelling reasons why the interaction picture cannot exist in the setting of a Fock space, at least for a scalar theory with a mass gap. They shed some light on the connection between the Fock representation with its characteristic number operator and go back to the publications [**DeDoRu66, DeDo67, Chai68**]. Strocchi has cast these results into a canonical form, ie purged of operator-algebraic argot, in his monograph [**Stro13**], which we shall follow in this section.

1.5.1. Fock space representations. We start with the usual creation and annihilation operators of a canonical free field, satisfying the CCR

$$[a(\mathbf{k}), a(\mathbf{k}')] = 0 = [a^\dagger(\mathbf{k}), a^\dagger(\mathbf{k}')] \ , \qquad [a(\mathbf{k}), a^\dagger(\mathbf{k}')] = (2\pi)^3\delta^{(3)}(\mathbf{k}-\mathbf{k}'). \tag{1.5.1}$$

and smooth them out with test functions $f \in \mathscr{S}(\mathbb{R}^3)$,

$$a(f) := \int \frac{d^3k}{(2\pi)^3} f^*(\mathbf{k}) a(\mathbf{k}), \qquad a^\dagger(f) := \int \frac{d^3k}{(2\pi)^3} f(\mathbf{k}) a^\dagger(\mathbf{k}). \tag{1.5.2}$$

Choosing an orthonormal Schwartz basis $f_j \in \mathscr{S}(\mathbb{R}^3)$ with respect to the inner product

$$(f_i, f_j) := \int \frac{d^3k}{(2\pi)^3} f_i^*(\mathbf{k}) f_j(\mathbf{k}) = \delta_{ij}, \tag{1.5.3}$$

the CCR (1.5.1) take the form

$$[a(f_i), a(f_j)] = 0 = [a^\dagger(f_i), a^\dagger(f_j)] \ , \qquad [a(f_i), a^\dagger(f_j)] = \delta_{ij}. \tag{1.5.4}$$

Then we obtain what we shall in the following call a *Heisenberg algebra*: if we set $a_j := a(f_j)$ and $a_j^* := a^\dagger(f_j)$, ie

$$[a_j, a_j] = 0 = [a_j^*, a_j^*] \ , \qquad [a_j, a_l^*] = \delta_{jl}. \tag{1.5.5}$$

The Heisenberg algebra is given by the polynomial algebra $\mathcal{A}_H := \langle a_j, a_j^* : j \in \mathbb{N}\rangle_{\mathbb{C}}$, generated by the creation and annihilation operators.

Given a vector Ψ_0 in a Hilbert space $\mathfrak{H}$ and a representation ϱ of the Heisenberg algebra, such that $\varrho(a_j)\Psi_0 = 0$ for all $j \in \mathbb{N}$, called *Fock representation*, it is clear that the *number operator* $N_\varrho := \sum_{j\geq 1} \varrho(a_j^*)\varrho(a_j)$ exists on the domain

$$\mathfrak{D}_0 := \varrho(\mathcal{A}_H)\Psi_0. \tag{1.5.6}$$

If the closure of $\mathfrak{D}_0$ yields $\mathfrak{H}$, ie if $\mathfrak{D}_0$ is dense in $\mathfrak{H}$, we say that the representation of the Heisenberg algebra is *cyclic* with respect to the vacuum Ψ_0. If $[C, \varrho(\mathcal{A}_H)] = 0$ implies $C = c1$ with $c \in \mathbb{C}$ for an operator C on $\mathfrak{H}$, the representation ϱ is called *irreducible.* Fock representations are always cyclic (by definition) and irreducible [**Stro13**]. We shall for convenience drop the symbol ϱ for the representation whenever there is no potential for confusion.

Vacuum & no-particle state. We shall now survey a small collection of assertions which urge us to draw the above mentioned conclusion about the interaction picture in Fock space. The first assertion is (tacitly) well-known among physicists.

PROPOSITION 1.4 (Number operator and vacuum [**Stro13**]). *Let $\{a_j, a_j^* : j \in \mathbb{N}\}$ be an irreducible representation of the Heisenberg algebra with a dense domain $\mathfrak{D}_0$ in a Hilbert space $\mathfrak{H}$. Then the following two conditions are equivalent.*

1. *The total number operator $N := \sum_{j\geq 1} a_j^* a_j$ has a nonnegative spectrum $\sigma(N)$ and exists in the sense that the strong limit*

$$s - \lim_{n\to\infty} e^{i\alpha \sum_{j=1}^n a_j^* a_j} = T(\alpha) \tag{1.5.7}$$

exists and defines a strongly continuous[13] *one-parameter group of unitary operators*[14].

2. *There exists a cyclic vector $\Psi_0 \in \mathfrak{H}$ such $a_j\Psi_0 = 0$ for all $j \in \mathbb{N}$.*

Proof. Let the first condition be given. First note that the CCR imply $T(\alpha)a_j = e^{-i\alpha}a_jT(\alpha)$ and $T(\alpha)a_j^\dagger = e^{i\alpha}a_j^\dagger T(\alpha)$. This entails $[T(2\pi), \mathcal{A}_H] = 0$ and thus $T(2\pi) = e^{i\theta}1$ on account of unitarity and irreducibility. Using the spectral representation of $T(\alpha)$, ie $T(\alpha) = \int_{\sigma(N)} e^{i\alpha\lambda} d\mathsf{E}(\lambda)$ we consider

$$0 = (T(2\pi) - e^{i\theta})^\dagger (T(2\pi) - e^{i\theta}) = \int_{\sigma(N)} |e^{i(2\pi\lambda-\theta)} - 1|^2 d\mathsf{E}(\lambda) \tag{1.5.8}$$

which means that the spectrum $\sigma(N)$ is discrete: $2\pi\lambda - \theta \in 2\pi\mathbb{Z}$. Pick $\lambda \in \sigma(N)$ with $\lambda > 0$ and let Ψ_λ be its eigenstate. Then

$$0 < \lambda||\Psi_\lambda||^2 = \langle\Psi_\lambda|N\Psi_\lambda\rangle = \sum_{j\geq1}\langle\Psi_\lambda|a_j^*a_j\Psi_\lambda\rangle = \sum_{j\geq1}||a_j\Psi_\lambda||^2 \tag{1.5.9}$$

entails that there is at least one j such that $a_j\Psi_\lambda \neq 0$. The CCR imply

$$T(\alpha)a_j\Psi_\lambda = e^{-i\alpha}a_jT(\alpha)\Psi_\lambda = e^{i\alpha(\lambda-1)}a_j\Psi_\lambda \tag{1.5.10}$$

and therefore $a_j\Psi_\lambda = c\Psi_{\lambda-1}$ with some $c \in \mathbb{C}$. Because the spectrum is bounded from below and nonnegative, there must be a state Ψ_0 such that $a_j\Psi_0 = 0$.

If the second condition is fulfilled, then it is clear that N exists on the sense subspace $\mathfrak{D}_0 = \mathcal{A}_H\Psi_0$ and so does the limit of the exponentiation. □

Hamiltonian chooses Fock representation. The next result is very interesting and pertinent to the interaction picture question.

Proposition 1.5 (Fock Hamiltonian [**Stro13**]). *Assume that for an irreducible representation $\{a_j, a_j^* : j \in \mathbb{N}\}$ of the Heisenberg algebra with dense domain $\mathfrak{D}_0$, there exists an operator*

$$H_0 = \sum_{j\geq1}\omega_j a_j^* a_j \ , \qquad \forall j: \ 0 < m \leq \omega_j \tag{1.5.11}$$

as generator of a strongly continuous one-parameter unitary group given by the strong limit

$$s - \lim_{n\to\infty} e^{i\alpha\sum_{j=1}^n \omega_j a_j^* a_j} = e^{i\alpha H_0} \tag{1.5.12}$$

such that the domain $\mathfrak{D}_0$ is stable under the action of this group. Then the Fock representation is selected by H_0.

Proof. $\sum_{j=1}^n \omega_j a_j^* a_j \geq m\sum_{j=1}^n a_j^* a_j$ tells us that the existence of H_0 implies the existence of the number operator N which by Proposition 1.4 selects the representation. □

This result relies on the existence of the mass gap, as the proof suggests. In fact, Proposition 1.5 can be formulated for a free scalar field with (formal) Hamiltonian

$$H_0 = \int \frac{d^3p}{(2\pi)^3}\ \omega(\mathbf{p})a^\dagger(\mathbf{p})a(\mathbf{p}) \tag{1.5.13}$$

with relativistic energy $\omega(\mathbf{p}) = \sqrt{\mathbf{p}^2 + m^2}$ of a particle with momentum $\mathbf{p} \in \mathbb{R}^3$ and rest mass $m > 0$. However, the case $m = 0$ is different: Proposition 1.5 does not hold and there is an infinite variety of inequivalent representations of the CCR, even with nonnegative energy [**BoHaSch63**].

[13]This means that the operators are *weakly* continuous in the parameter α, clearly a very confusing convention in functional analysis! See Appendix Section A.1.

[14]Stone's theorem then guarantees that N exists as the generator of $T(\alpha)$.

We shall therefore stick to the mass gap case. The Hamiltonian (1.5.13) can be defined as the operator acting on the one-particle state $a^\dagger(f)\Psi_0$ according to

$$H_0 a^\dagger(f)\Psi_0 = [H_0, a^\dagger(f)]\Psi_0 = a^\dagger(\omega f)\Psi_0, \tag{1.5.14}$$

in which $(\omega f)(\mathbf{k}) = \omega(\mathbf{k})f(\mathbf{k})$ is a perfect Schwartz function, cf.(1.5.2). This definition coheres with the formal Hamiltonian (1.5.13) and the CCR (1.5.1). Using the condition $H_0\Psi_0 = 0$, one can then easily compute

$$H_0 a^\dagger(f_1)...a^\dagger(f_n)\Psi_0 = [H_0, a^\dagger(f_1)...a^\dagger(f_n)]\Psi_0 \tag{1.5.15}$$

by using the commutator property $[A, BC] = B[A, C] + [A, B]C$. The matrix element of H_0 with respect to the one-particle state is

$$\langle a^\dagger(f)\Psi_0 | H_0 a^\dagger(f)\Psi_0\rangle = \langle a^\dagger(f)\Psi_0 | a^\dagger(\omega f)\Psi_0\rangle = (f, \omega f). \tag{1.5.16}$$

One can now choose f such that the rhs of (1.5.16) cannot be distinguished by measurement from the correct relativistic energy of a free particle with some fixed momentum.

In the same way, one introduces the momentum operator,

$$P^j = \int \frac{d^3p}{(2\pi)^3}\, p^j a^\dagger(\mathbf{p})a(\mathbf{p}) \tag{1.5.17}$$

which is given through $[P^j, a^\dagger(f)] = a^\dagger(p^j f)$ and completes the translation subalgebra of the Poincaré group. The creators and annihilators then transform according to

$$e^{ib\cdot P}a(f)e^{-ib\cdot P} = a(e^{ib\cdot p}f)\ , \qquad e^{ib\cdot P}a^\dagger(f)e^{-ib\cdot P} = a^\dagger(e^{ib\cdot p}f). \tag{1.5.18}$$

For for pure time translations, this takes the form

$$e^{iH_0t}a(f)e^{-iH_0t} = a(e^{i\omega t}f)\ , \qquad e^{iH_0t}a^\dagger(f)e^{-iH_0t} = a^\dagger(e^{i\omega t}f). \tag{1.5.19}$$

The Lorentz group is implemented similiarly. Finally, the canonical free field $\varphi_0(t, f)$ and its conjugate momentum field are then given by

$$\varphi_0(t,f) = \frac{1}{\sqrt{2}}[a(\omega^{-\frac{1}{2}}e^{i\omega t}f) + a^\dagger(\omega^{-\frac{1}{2}}e^{i\omega t}f)],\ \ \pi_0(t,f) = \frac{i}{\sqrt{2}}[a^\dagger(\omega^{\frac{1}{2}}e^{i\omega t}f) - a(\omega^{\frac{1}{2}}e^{i\omega t}f)] \tag{1.5.20}$$

and satisfy the CCR

$$[\varphi_0(t,f), \varphi_0(t,g)] = 0 = [\pi_0(t,f), \pi_0(t,g)]\ , \qquad [\varphi_0(t,f), \pi_0(t,g)] = i(f,g). \tag{1.5.21}$$

This is what we call the Fock representation of a free field. One distinguishes the translation-invariant state called vacuum from the no-particle state. Because of $P^\mu\Psi_0 = 0 = N\Psi_0$, the intuition is that both coincide. Now consider

Proposition 1.6 (Fock representation [**Stro13**]). *All Fock representations are unitarily equivalent. The vacuum is unique, ie the only translation-invariant state, and coincides with the no-particle state.*

Proof. The isomorphism $V : \varrho(\mathcal{A}_H)\Psi_0 \to \varrho'(\mathcal{A}_H)\Psi_0'$ is densely defined and preserves the scalar product. Let Ψ_0 be the no-particle state, ie $N\Psi_0 = 0$ and let $\Psi \neq \Psi_0$ also have this property. Then $a_j\Psi = 0$ and hence $\langle\Psi|A\Psi_0\rangle = 0$ if $A \in \mathcal{A}_H^* := \langle a_j^* : j \in \mathbb{N}\rangle_{\mathbb{C}}$. But because $\mathcal{A}_H\Psi_0 = \mathcal{A}_H^*\Psi_0$ is dense, we find $\Psi = 0$, ie the state Ψ_0 is the only state annihilated by the a_j's. Now note that N commutes with the space translation operator $U(\mathbf{b}) = \exp(-i\mathbf{b}\cdot\mathbf{P})$ by definition of P^j in (1.5.17). Let $\Phi \neq \Psi_0$ be translation invariant. Then,

$$U(\mathbf{b})N\Phi = NU(\mathbf{b})\Phi = N\Phi = \sum_{n\geq 0} N\Phi_n \tag{1.5.22}$$

where $\Phi = \sum_{n\geq 0}\Phi_n$ is the decomposition into the n-particle subspace components. This implies $U(\mathbf{b})\Phi_n = \Phi_n$, ie a contradiction because the n-particle state is not translation invariant. $\square$

1.5.2. Interaction picture. Now, in the light of the above results, the implementation of the interaction picture poses the following problem. The intertwining operator for the interaction picture[15],

$$V(t) = e^{iH_0 t} e^{-iHt} \tag{1.5.23}$$

demands the separate existence of a free Hamiltonian H_0 which by Propositions 1.5 and 1.6 selects a Fock representation with a unique Poincaré-invariant vacuum. The interacting Hamiltonian H is yet another, an additional time translation generator that needs to be implemented. The time evolution operator for interaction picture states, given by

$$\mathsf{U}(t,s) = V(t)V(s)^\dagger = e^{iH_0 t} e^{-iH(t-s)} e^{-iH_0 s} \tag{1.5.24}$$

clearly shows that H and H_0 are supposed to act in the same Hilbert space, ie the Fock space. The problem is now that even in relatively simple models studied so far, this split into two well-defined self-adjoint operators $H = H_0 + H_{int}$ has not been possible: either the sum is well-defined or only one part of it, not both. The requirement of some form of renormalisation then always leads to non-Fock representations, which by Proposition 1.6 cannot be unitarily equivalent to the Fock representation of a free field (see [**Stro13**], pp. 40).

Another compelling argument put forward by Wightman in [**Wi67**] against the existence of the interaction picture uses translation invariance to show that the Schrödinger picture integral

$$H_{int} = \int d^3x \; \mathscr{H}_{int}(\mathbf{x}), \tag{1.5.25}$$

ie the interacting part of the Hamiltonian, does not make sense. Because the interaction picture Hamiltonian is given by

$$H_I(t) := e^{iH_0 t} H_{int} e^{-iH_0 t} = \int d^3x \; e^{iH_0 t} \mathscr{H}_{int}(\mathbf{x}) e^{-iH_0 t} =: \int d^3x \; \mathscr{H}_I(t,\mathbf{x}) \tag{1.5.26}$$

we will see there is no doubt that something is wrong. It is a head-on attack on the interaction picture and the reasoning is straightforward. Here is Wightman's argument: we first compute

$$\begin{aligned} ||H_{int}\Psi_0||^2 &= \int d^3x \int d^3y \; \langle \Psi_0 | \mathscr{H}_{int}(\mathbf{x}) \mathscr{H}_{int}(\mathbf{y}) \Psi_0 \rangle \\ &= \int d^3x \int d^3y \; \langle \Psi_0 | \mathscr{H}_{int}(0) \mathscr{H}_{int}(\mathbf{y}) \Psi_0 \rangle . \end{aligned} \tag{1.5.27}$$

This integral over $\mathbf{x} \in \mathbb{R}^3$ diverges unless $\mathscr{H}_{int}(\mathbf{y})\Psi_0 = 0$, at least according to Wightman in [**Wi67**]; although one does not have to follow him here, the integrand must vanish in any case. And because the evolution operators in (1.5.26) can be inserted without changing the norm, we find that $||H_I(t)\Psi_0|| = 0$ is the only sensible and acceptable outcome.

But because the Hamiltonian $\mathscr{H}_I(t,\mathbf{x})$ is made up of free interaction picture fields which always have terms with the right combination of creators and annihilators for the vacuum expectation value not to vanish, we can see that the interaction picture Hamiltonian does not exist in the way the canonical formalism desires it to. However, contrary to Wightman's conclusion, one can interprete (1.5.27) in such a way as to say

$$\langle \Psi_0 | \mathscr{H}_I(t,0) \mathscr{H}_I(t,\mathbf{y}) \Psi_0 \rangle = \langle \Psi_0 | \mathscr{H}_{int}(0) \mathscr{H}_{int}(\mathbf{y}) \Psi_0 \rangle = 0 \tag{1.5.28}$$

just means that the states $\mathscr{H}_I(t,0)\Psi_0$ and $\mathscr{H}_I(t,\mathbf{y})\Psi_0$ have no overlap. But still, if these Hamiltonians are composed of monomials of free interaction picture fields, this is not acceptable.

[15]See Section 3.1 for a review of this part of the canonical formalism.

1.6. Canonical (anti)commutation relations and no-interaction theorems

It is important to note that the proof of Haag's theorem does not require any of the fields to obey the canonical (anti)commutation relations (CCR/CAR) *explicitly* which, for the spatially smeared scalar field operators

$$\phi(t,f) = \int d^n x \; f(\mathbf{x})\phi(t,\mathbf{x}) \;, \qquad \pi(t,g) = \int d^n x \; g(\mathbf{x})\pi(t,\mathbf{x}) \qquad f,g \in \mathscr{S}(\mathbb{R}^n) \tag{1.6.1}$$

take the form

$$[\phi(t,f),\phi(t,g)] = 0 = [\pi(t,f),\pi(t,g)] \;, \qquad [\phi(t,f),\pi(t,g)] = i(f,g), \tag{1.6.2}$$

where $n \geq 1$ is the dimension of space[16]. This condition is, however, *implied* in a trivial way for the following reason: if (1.6.1) are free fields and the unitary map V also transforms the conjugate momentum field, then they of course satify the CCR (1.6.2). Applying the unitary transformation that connects both field theories to (1.6.2) shows that the other fields inevitably obey the CCR, too. Hence both theories are unitarily equivalent representations of the CCR.

This uncorks the question whether the CCR (1.6.2) are somehow related to the triviality result entailed by Haag's theorem. Can truely interacting fields be *any* representation of the these commutation relations? We shall see now that here, in contrast to Haag's theorem, the dimension of spacetime is decisive.

1.6.1. Anticommutation relations and triviality. Under some regularity conditions, Powers answered this question for fermion fields in [**Pow67**]. For a Dirac field $\psi(t,f)$ and its canonical conjugate field $\psi(t,f)^\dagger = \psi^\dagger(t,f^*)$ the canonical anticommutation relations (CAR) take the form

$$\{\psi(t,f),\psi(t,g)^\dagger\} = 0 = \{\psi(t,f)^\dagger,\psi(t,g)^\dagger\} \;, \qquad \{\psi(t,f),\psi(t,g)^\dagger\} = i(f,g). \tag{1.6.3}$$

Powers' result is now that these relations imply triviality in space dimension $n \geq 2$.

Theorem 1.7 (Powers' theorem). *Let $\psi(t,f)$ be a local relativistic Fermi field in the sense of Wightman's framework in $d = n+1 \geq 3$ spacetime dimensions fulfilling the CAR (1.6.3) and acting together with its adjoint $\psi(t,f)^\dagger$ in a Hilbert space $\mathfrak{H}$ with vacuum state Ω_0.*

Assume that they form an irreducible set of operators at one fixed instant and that there is a unitary transformation $U(t)$ such that

$$\psi(t,f) = U(t)\psi(0,f)U(t)^\dagger \tag{1.6.4}$$

for all times $t \in \mathbb{R}$ and that the limits

$$\begin{gathered} \lim_{t\to 0}\frac{1}{t}[\psi(t,f)-\psi(0,f)]\Omega_0 = \partial_t\psi(0,f)\Omega_0 \;, \\ \lim_{t\to 0}\frac{1}{t}[\psi(t,f)-\psi(0,f)]\psi(0,f)\Omega_0 = \partial_t\psi(0,f)\psi(0,f)\Omega_0 \end{gathered} \tag{1.6.5}$$

and the corresponding ones for the adjoint exist in the norm for all test functions $f \in \mathscr{S}(\mathbb{R}^n)$. Then $\psi(t,f)$ is a free field in the sense that it satisfies a linear differential equation which is first order in time.

Proof. See [**Pow67**]. Powers has developed and employed techniques that Baumann used in his proof of Theorem 1.8, which is sketched in Appendix Section B.1. □

[16] The smeared fields in (1.5.20) and (1.6.1) differ: f in (1.6.1) is the Fourier transform of f in (1.5.20), but never mind.

A few comments are in order. Powers did *not* prove that the fields satisfy the Dirac equation. Instead he found that the conditions imposed on the Fermi field are so restrictive that there exist operators T_1 and T_2, linear and antilinear on $\mathscr{S}(\mathbb{R}^n)$, respectively, such that

$$\partial_t \psi(t, f) = \psi(t, T_1 f) + \psi(t, T_2 f)^\dagger. \tag{1.6.6}$$

These operators may be realised in such a way that T_1 is the spatial part of the Dirac operator (stripped of the zeroth γ-matrix) and $T_2 = 0$. They do not have to take this specific form, though. This is the precise sense in which the fields are free.

Surely, if an interacting Dirac field obeys a differential equation in whatever sense, it should not be of the form (1.6.6) because this implies in particular that there are no other fields to interact with. The assumption that these fields form an irreducible set of operators is, up to some mathematical subleties, equivalent to their capability of generating a dense subspace[17]. In other words, Powers' theorem is an assertion about a quantum field theory with fermions only. The theories that spring to mind are of the Fermi theory type with 'four-fermion' interactions like

$$(\overline{\psi}\gamma_\mu\psi)(\overline{\psi}\gamma^\mu\psi) \qquad \text{('Thirring model')} \tag{1.6.7}$$

which are all together non-renormalisable and hence unphysical, albeit of some value as effective field theories. Consequently, if we from a physical point of view require renormalisability, fermions cannot directly interact with each other and should thus be free. Therefore, Powers' theorem makes perfect sense.

1.6.2. Commutation relations and triviality. Baumann investigated the case of bosonic fields but could not find the exact analogue of Powers' result [**Bau87**]. Yet he proved that in space dimension $n > 3$ only free theories can satisfy the CCR (1.6.2) and suggests that the theories $(\phi^4)_4$ and $(\phi^6)_3$ may fulfill these relations, whereas for $(\phi^4)_2$, he found 'no restrictions', in agreement with Glimm and Jaffe's results.

In fact, the only interacting theories so far found to satisfy the CCR are superrenormalisable field theories like $P(\phi)_2$, the sine-Gordon model $\cos(\phi)_2$ and the exponential interaction $\exp(\phi)_2$ [**GliJaf81**].

Baumann's provisions fill a rather long list, to be looked up in [**Bau87**] by the interested reader. To present them here would coerce us to explain earlier results by Herbst and Fröhlich, including some argot (see the references in [**Bau87**]). We shall not embark on this here and content ourselves with the main assertion and a sketch of the proof relegated to the appendix.

THEOREM 1.8 (Baumann). *Let $n \geq 4$ be the space dimension and $\varphi(t, \cdot)$ a scalar field with conjugate momentum field $\pi(t, \cdot) = \partial_t \varphi(t, \cdot)$ such that the CCR (1.6.2) are obeyed and assume furthermore that $\dot{\pi}(t, \cdot) := \partial_t \pi(t, \cdot)$ exists. Then, if $\varphi(t, \cdot)$ has a vanishing vacuum expectation value and the provisions listed in the introduction of* [**Bau87**] *are satified, one has*

$$\dot{\pi}(t, f) - \varphi(t, \Delta f) + m^2 \varphi(t, f) = 0 \tag{1.6.8}$$

for all $f \in \mathscr{D}(\mathbb{R}^n)$ and a parameter $m^2 > 0$.

PROOF. See Appendix Section B.1. □

Notice that $\pi(t, \cdot) = \partial_t \varphi(t, \cdot)$ is not correct if the Lagrangian has an interaction term with a first time derivative, as we have in the case of a renormalised field theory incurred by the counterterms. This line of argument is heuristic, but we have no cause to believe $\pi = \dot{\varphi}$ either.

When Baumann scrutinised the case in which both coexisting bosons and fermions satisfy the CCR and CAR, respectively, he again found that for $n > 3$ space dimensions any theory with these relations must necessarily be free while for $n = 3$ it was impossible to say [**Bau88**]. We remind the reader that no renormalisable and non-renormalisable models have been constructed

[17] See Section 2.2 on this issue.

so far (see Introduction): it was merely Baumann's working assumption that these interacting theories exist and that their interaction terms have no derivative coupling.

Unfortunately, space dimension $n = 3$ remained defiant. Baumann mentions that the unpublished proof for $n \geq 3$ by Sinha, then ostensibly a PhD student of Emch's [**SinEm69**], had weaker assumptions that somehow did not appeal to him. Emch presents Sinha's version without proof in [**Em09**]. From what we can tell by comparing both results, Sinha's provisions are weaker as he uses the Weyl representation of the CCR instead of the CCR themselves. We shall merely quote[18] the Weyl CCR case from [**Em09**].

THEOREM 1.9 (Sinha). *Let $n \geq 3$ be the dimension of space and $\phi(t, f)$ a sharp-time local scalar Wightman field with canonical conjugate $\pi(t, f)$ that generate a representation of the Weyl CCR at $t = 0$, ie*

$$U(f) = e^{i\varphi(0,f)}, \; V(g) = e^{i\pi(0,g)}, \qquad U(f)V(g) = e^{i(f,g)}V(g)U(f) \tag{1.6.9}$$

where $U(f)U(g) = U(f+g)$ and $V(f)V(g) = V(f+g)$ for all $f, g \in \mathscr{S}(\mathbb{R}^n)$ such that the families $\lambda \mapsto U(\lambda f)$ and $\lambda \mapsto V(\lambda f)$ are weakly continuous. Assume further that $\partial_t \pi(t, f)$ exists and that the Weyl unitaries $U(f), V(g)$ are irreducible with common dense domain $\mathfrak{D}$ which is stable under the algebra of their generators. Then there exists a linear operator $T \colon \mathscr{S}(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ and a distribution $c \in \mathscr{S}'(\mathbb{R}^n)$ such that

$$\partial_t^2 \phi(t, f) = \phi(t, Tf) + c(f) \tag{1.6.10}$$

and $\phi(t, \cdot), \pi(t, \cdot)$ fulfill the CCR (1.6.2).

These no-interaction theorems by Powers, Baumann and Sinha suggest that renormalisable and fully fledged interacting field theories will most likely neither satisfy the CCR (1.6.2) nor the CAR (1.6.3). In other words, for a quantum field theory, there is no analogous representation issue as in quantum mechanics: we had better not seek a unitarily equivalent representation of the Fock representation since that will in all likelyhood be a free theory.

For theories of only one type of field, ie fermions or scalar bosons, these results can be viewed as a variant of Haag's theorem in the sense that Dyson's matrix cannot exist: if a quantum field and its canonical momentum field are unitarily equivalent to those of a free theory at one fixed instant $t = t_0$, then it satisfies the canonical (anti)commutation rules at that time. By the foregoing theorems, the field can only be free in spacetime dimensions $d \geq 5$ for bosons and $d \geq 3$ for fermions. If one assumes the Weyl form of the CCR, boson fields are free for spacetime dimensions $d \geq 4$.

The interesting issue here is the spacetime dependence: we know that superrenormalisable theories of the type $P(\varphi)_2$ conform with the CCR [**GliJaf81**], whereas (probably) renormalisable and non-renormalisable ones do not. Notice that these results do not mean that interacting theories cannot exist in higher-dimensional spacetimes. Powers' and Baumann's results merely inform us that interactions are incompatible with the CCR/CAR there.

1.6.3. CCR/CAR and the Heisenberg uncertainty principle. In Subsection 1.1.2 we have already discussed where the (anti)commutation relations came from: the Heisenberg uncertainty principle in quantum mechanics. Although free fields enjoy these relations by construction, we do not need them for interacting theories.

One reason is philosophical in nature: there is no position operator in QFT and the relation of the CCR/CAR to the Heisenberg uncertainty principle is obscure, at least in our mind and to the best of our knowledge.

[18]The only source containing the proof seems to be Sinha's PhD thesis, only existing in print at the library of the University of Florida, whose staff did not reply to the our email. We did not insist.

The other is practical: any particles measured in scattering experiments are detected *after* the scattering event when they are deemed free. The measuring apparatus cannot be placed within the interaction vertex[19].

Besides, concrete computations leading to numbers that can be measured in experiments are always carried out with free fields. It is through these fields that the constant $\hbar$ enters the theory. Unfortunately, none of the authors, ie Baumann, Powers, Strocchi and Wightman, who proposed that we abandon the CCR/CAR for a general interacting QFT, touched upon this important issue. This also goes for the next author whose pertinent results we shall discuss in brief.

1.6.4. Lopuszanski's contribution. We finally mention Lopuszanski's results because of their relevance to the above no-interaction theorems. The aim of his work was to extensively characterise free scalar fields in order to figure out what properties interacting fields can by exclusion not have.

We briefly review some of his results published in [**Lo61**]. We start with the Yang-Feldman representation of a massive scalar interacting Heisenberg field

$$\phi(t,\mathbf{x}) = \phi_{\mathrm{in}}(t,\mathbf{x}) - \int d^4y\ \Delta_{\mathrm{ret}}(x-y)j(y), \tag{1.6.11}$$

where $\Delta_{\mathrm{ret}}(x-y)$ is the retarded Green's function of the Klein-Gordon operator and $j(y)$ the interaction term from the equation of motion, possibly containing other fields. The field $\phi_{\mathrm{in}}(t,\mathbf{x})$ describes free incoming and asymptotic bosons. Because the retarded Green's function vanishes in the limit $t = x^0 \to \pm\infty$, the Heisenberg field $\phi(t,\mathbf{x})$ converges to the asymptotic field by the way it is defined[20].

Lopuszanski's results interest us here because they make plausible that another path to the construction of the S-matrix which circumvents the interaction picture and hence makes no use of Dyson's matrix, is equally haunted by triviality if the provisions are too strong: the representation of the S-matrix in terms of Heisenberg fields due to Yang and Feldman [**YaFe50**].

The first assumption is that the interaction current takes the form

$$j(y) = \frac{g}{3!}\phi(y)^3. \tag{1.6.12}$$

This is already problematic because powers of fields are ill-defined and must be Wick-ordered as we will discuss in the next chapter, Subsection 3.3.1 (see also [**StreatWi00**], p.168).

The assumption that the Heisenberg field can be representated as a Fourier mode expansion

$$\phi(t,\mathbf{x}) = \int \frac{d^3q}{(2\pi)^3}\frac{1}{\sqrt{2E_q}}\ [e^{-iq\cdot x}a(t,q) + e^{iq\cdot x}a^\dagger(t,q)], \tag{1.6.13}$$

with $E_q := \sqrt{\mathbf{q}^2+m^2}$ is already a bit strong for an interacting field. These mode operators are assumed to satisfy[21]

$$\langle 0|[a(t,q), a^\dagger(t,q')]|0\rangle = Z^{-1}\delta^{(3)}(\mathbf{q}-\mathbf{q}') \tag{1.6.14}$$

where all other commutators vanish and Z^{-1} is (by Lopuszanski's assumption) the *finite* inverse wavefunction renormalisation. This amounts to demanding that the Heisenberg field obey the CCR, albeit including a peculiar factor. The only way this theory has a chance to differ from a free one lies in the time dependence of the mode operators.

[19]When the wire chamber detects a *Townsend avalanche* kicked off by, say, ionised argon, everything is over already: whatever the interaction vertex's size (far below atomic scale), what comes out of it can only be thought of as free.

[20]The Yang-Feldman equation (1.6.11) says that the outgoing field is identical to the incoming one.

[21]We adopt Lopuszanski's notation $|0\rangle$ for the vacuum.

The idea that ϕ converges to the free field ϕ_{in} as we go back to the remote past means that these operators converge in some sense to those of the incoming field, ie

$$a(t,q) \;\to\; a_{\text{in}}(q) \;, \qquad a^\dagger(t,q) \;\to\; a^\dagger_{\text{in}}(q) \qquad \text{as } t \to -\infty \;. \tag{1.6.15}$$

This entails $Z = 1$ because in the limit, the lhs of (1.6.14) goes over to the commutator of the incoming mode operators which require $Z = 1$, as this object is necessarily time-independent. Because, so he argues, the Källen-Lehmann spectral representation must satisfy

$$Z^{-1} = \int d\mu^2\ \rho(\mu^2) = \int d\mu^2\ [\ \delta(\mu^2 - m^2) + \sigma(\mu^2)\ ] = 1 + \int d\mu^2\ \sigma(\mu^2) \tag{1.6.16}$$

the conclusion is $\sigma(\mu^2) = 0$, that is, $\phi(t, \mathbf{x})$ is trivial. This is Lopuszanski's first no-interaction result which tells us that the CCR (1.6.14) had better not be fulfilled.

We feel strongly obliged to critique (1.6.16), but will defer a discussion of this issue and the wave renormalisation constant to Section 1.7). Let us now have a look at Lopuszanski's main theorem.

Claim 1.10 (Lopuszanski). *Assume $\phi(x)|0\rangle = \phi_{in}(x)|0\rangle$. Then follows that $\phi(x) = \phi_{in}(x)$ and the theory is trivial.*

Proof. The proof is elementary: first of all, the assumption tells us $j(y)|0\rangle = 0$. This means the interaction term is incapable of polarising the vacuum. This seems questionable for (1.6.12), to say the least. However, because of causality, we have

$$[\phi(t, \mathbf{x}), j(t, \mathbf{y})] = 0 \tag{1.6.17}$$

and thus $0 = \phi(t, \mathbf{x}) j(t, \mathbf{y})|0\rangle = j(t, \mathbf{y})\phi(t, \mathbf{x})|0\rangle = j(t, \mathbf{y})\phi_{\text{in}}(t, \mathbf{x})|0\rangle$. This is nothing but

$$0 = \int \frac{d^3q}{(2\pi)^3} \frac{1}{\sqrt{2E_q}}\ e^{-ix\cdot q} j(t, \mathbf{y}) a^\dagger_{\text{in}}(q)|0\rangle \tag{1.6.18}$$

and therefore $j(t, \mathbf{y})|q\rangle = 0$. By induction one gets $j(y)|q_1, ..., q_n\rangle = 0$. Finally, on account of the assumption that these states span a dense subspace, the result is $j(y) = 0$. □

Lopuszanski's assumption $\phi(x)|0\rangle = \phi_{\text{in}}(x)|0\rangle$ was probably inspired by the idea that applying the interacting field only once should create *one* single particle. Because there is no other particle, the particle created by the free field cannot be distinguished from a single particle minted by the interacting field.

But this assumption already sneaks in that there is no vacuum polarisation: once the particle has been created, it will be there on its own and there will be nothing for it to interact with. No cloud around it, hence no interaction.

We are thus informed that this assumption is fallacious for an interacting field. Lopuszanski himself concludes that a reasonable interacting field should not be of the form (1.6.13) and should also not satisfy the CCR (1.6.14) [**Lo61**].

1.6.5. Conclusion about the CCR/CAR. In 1964 and hence prior to the publication of the above no-interaction theorems, Streater and Wightman wrote in their book [**StreatWi00**] that they do not exclude the CCR for interacting fields in general, but contend that the hints they have from examples leave them in no doubt that singular behaviour is to be expected for sharp-time fields, even after being smoothed out in space.

Consequently, it may in such cases be difficult to give the CCR in (1.6.2) a meaning. “Thus, one is reluctant to accept canonical commutation relations as an indispensible requirement on a field theory.” ([**StreatWi00**], p.101).

Although Wightman encouraged Baumann to work on the CCR/CAR question (see acknowledgements in [**Bau87, Bau88**]), he did obviously not deem the results important enough to update these remarks in the latest edition [**StreatWi00**] of the year 2000.

So we conclude that although some superrenormalisable theories have been found to conform with the CCR, renormalisable and non-renormalisable theories cannot be expected to have this feature. However, if an interacting field theory fulfills what is known as the *asymptotic condition*, then it may obey these relations at least *asymptotically.*

The idea that interacting fields obey some form of the CCR or CAR is generally not discussed and strictly speaking not claimed to be true in physics. In fact, when asked, many practicing physicists would have to first think of where the idea came from to produce the Heisenberg uncertainty principle as an answer.

Yet the CCR/CAR represent a property constitutive for the quantisation of free fields. Therefore, it is somehow tacitly taken for granted for 'all' fields when a classical theory is 'quantised'. But as a property, it is, in actual fact, only used for free fields to compute the Feynman propagator which is in turn needed for perturbation theory in the interaction picture.

We shall see in Section 3.2 that the CCR question cannot be answered within the canonical formalism of perturbation theory in its current form.

On the grounds that these relations are simply of no practical relevance for interacting fields and give rise to philosophical questions only, we are of the opinion that for the time being, one can easily get on without them and hence need not be disturbed by Powers' and Baumann's results.

Our impression is that these triviality results are generally unknown to practising physicists[22] and given that *those in the know about them are only mathematical physicists of a special creed*, namely axiomatic and algebraic quantum field theorists, this situation is likely to stay that way.

Let us welcome Powers' and Baumann's theorems as another piece of information about interacting field theories that make us realise how little we know about them and how much our imagination is influenced by being only familiar with free fields. But because a large class of superrenormalisable theories conform with the CCR, the question as to why (or why not) still lingers on.

1.7. Wave-function renormalisation constant

As alluded to above, other mathematical physicists have also expressed their doubts about the CCR for interacting fields. In his recent monograph, Strocchi comes to the same devastating conclusion about the CCR for interacting fields, he writes "...canonical quantization cannot be used as a rigorous method for quantizing relativistic interacting fields..."([**Stro13**], p.51).

1.7.1. CCR/CAR and wave-function renormalisation. He mentions the no-interaction results of Powers and Baumann and expresses his opinion that the singular behaviour of interacting sharp-time fields is the root of all evil. His main argument is to say that the CCR for the renormalised and hence interacting field

$$[\varphi_r(t,\mathbf{x}), \dot{\varphi}_r(t,\mathbf{x})] = iZ^{-1}\delta^{(3)}(\mathbf{x}-\mathbf{y}) \tag{1.7.1}$$

make no sense because - according to Strocchi - the wave renormalisation Z vanishes and lets the rhs diverge. Before we comment on this, let us quickly review where this form of the CCR comes from and that, in fact, the canonical Lagrangian formalism protects itself from attacks like this one. Both (1.6.14) and (1.7.1) are the result of assuming

$$[\varphi(t,\mathbf{x}), \dot{\varphi}(t,\mathbf{x})] = [\varphi(t,\mathbf{x}), \pi(t,\mathbf{x})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y}) \tag{1.7.2}$$

for the bare fields[23] $\varphi(x)$ and $\pi(x)$, where the canonical momentum field is found by differentiating the bare Lagrangian with respect to $\dot{\varphi} = \partial_t\varphi$ and simply yields $\pi(x) = \dot{\varphi}(x)$. One then

[22]Otherwise there should at least be a short remark about it in every lecture/textbook of QFT when the CCR/CAR are introduced. Since free fields satisfy them, no restrictions follow and the lecture can blithely be carried on.

[23]Notice that the bare field is not to be confused with a free field.

takes the renormalised field, $\varphi_r := Z^{-1/2}\varphi$ and its first time derivative $\dot\varphi_r = Z^{-1/2}\dot\varphi$ to find for the renormalised field

$$[\varphi_r(t,\mathbf{x}), \dot\varphi_r(t,\mathbf{x})] = Z^{-1}[\varphi(t,\mathbf{x}), \dot\varphi(t,\mathbf{x})] = iZ^{-1}\delta^{(3)}(\mathbf{x}-\mathbf{y}). \tag{1.7.3}$$

But this is *not* the 'proper' CCR, ie what comes out if we strictly follow the rules of the canonical formalism: the 'correct' renormalised conjugate momentum is given by

$$\pi_r = \frac{\partial \mathscr{L}}{\partial \dot\varphi_r} = Z\dot\varphi_r = ZZ^{-1/2}\dot\varphi = Z^{1/2}\pi. \tag{1.7.4}$$

where one has to make use of the Lagrangian of the renormalised field, given by

$$\mathscr{L} = \frac{1}{2}Z(\partial\varphi_r)^2 - \frac{1}{2}m_r^2 Z_m \varphi_r^2 - \frac{g_r}{4!}Z_g\varphi_r^4, \tag{1.7.5}$$

which we will discuss extensively in Section 3.4. Then, with this result, we find by (1.7.4)

$$[\varphi_r(t,\mathbf{x}), \pi_r(t,\mathbf{x})] = [Z^{-1/2}\varphi(t,\mathbf{x}), Z^{1/2}\pi(t,\mathbf{x})] = i\delta^{(3)}(\mathbf{x}-\mathbf{y}), \tag{1.7.6}$$

which is the 'canonically correct' CCR of the renormalised field, completely free of ailments, seemingly.

1.7.2. Lopuszanski's argument. We resume the discussion on the CCR question in Subsection 1.6.4, where we presented Lopuszanski's reasoning. To derive (1.6.16), we first consider the commutator function of the free field $\phi_0(t,\mathbf{x})$,

$$D(t-s,\mathbf{x}-\mathbf{y};m^2) := [\phi_0(t,\mathbf{x}), \phi_0(s,\mathbf{y})] = \int \frac{d^4q}{(2\pi)^3}\ \delta_+(q^2-m^2)[e^{-iq\cdot(x-y)} - e^{+iq\cdot(x-y)}], \tag{1.7.7}$$

then take the Källen-Lehmann representation

$$\langle\Omega|[\phi(t,\mathbf{x}), \phi(s,\mathbf{y})]\Omega\rangle = \int d\mu^2\ \rho(\mu^2) D(t-s,\mathbf{x}-\mathbf{y};\mu^2) \tag{1.7.8}$$

of the interacting (renormalised) field's commutator and differentiate it with respect to s to get

$$\langle\Omega|[\phi(t,\mathbf{x}), \dot\phi(s,\mathbf{y})]\Omega\rangle = -\int d\mu^2\ \rho(\mu^2)\ \partial_t D(t-s,\mathbf{x}-\mathbf{y};\mu^2) \tag{1.7.9}$$

where the identity $\partial_s D(t-s,\cdot) = -\partial_t D(t-s,\cdot)$ for the integrand is obvious. On account of

$$\partial_t D(0,\mathbf{x}-\mathbf{y};\mu^2) = -i\delta^{(3)}(\mathbf{x}-\mathbf{y}) \tag{1.7.10}$$

we find that in the limit $s \to t$ the Källen-Lehmann representation in (1.7.9) goes over to

$$\langle\Omega|[\phi(t,\mathbf{x}), \dot\phi(t,\mathbf{y})]\Omega\rangle = i\int d\mu^2\ \rho(\mu^2)\ \delta^{(3)}(\mathbf{x}-\mathbf{y}) = i[\ 1 + \int d\mu^2\ \sigma(\mu^2)\ ]\delta^{(3)}(\mathbf{x}-\mathbf{y}), \tag{1.7.11}$$

and $Z^{-1} = 1 + \int d\mu^2\ \sigma(\mu^2)$ is the conclusion if the field $\varphi(x)$ obeys the CCR (1.7.1). The canonical assertion $0 \le Z \le 1$ and $Z = 1$ for free fields is then an easy consequence.

1.7.3. Wave-function renormalisation. Strocchi contends that $Z \to 0$ upon removal of the cutoff because the spectral integral

$$Z^{-1} = \int d\mu^2\ \rho(\mu^2) = 1 + \int d\mu^2\ \sigma(\mu^2) \tag{1.7.12}$$

diverges in four spacetime dimensions "as a consequence of general non-perturbative arguments" ([**Stro13**], p.51), at which point he cites Powers' and Baumann's papers [**Bau87, Pow67**].

From our understanding of those nonperturbative arguments, we assume the idea behind Strocchi's statement is something like this:

(1) free fields satisfy the CCR and besides $\int d\mu^2\ \rho(\mu^2) = 1$ is uncontroversial. Hence (1.7.11) makes total sense for free fields, whereas

(2) interacting fields do not satisfy the CCR (in dimensions $d \geq 5$), ie something must go wrong.
(3) Conclusio: (1.7.11) must diverge.

Whatever the author had in mind, we do not find this convincing. Apart from the fact that Baumann's results do strictly speaking *not pertain* to four spacetime dimensions, the problem may rather lie in the provision that a sharp-time Wightman field and its first derivative with respect to time exist.

In four spacetime dimensions, these dubious two objects possibly only exist in the trivial free case. One therefore cannot use them to conclude that the spectral integral (1.7.12) diverges.

However, on page 106 in [**Stro13**], Strocchi discusses the wave-function renormalisation constant Z_Λ with a UV cutoff $\Lambda > 0$ for the Dirac field ψ of the *derivative coupling model*

$$\mathscr{L}_{DC} = \frac{1}{2}[\partial_\mu\varphi\partial^\mu\varphi - m^2\varphi^2] + \overline{\psi}[i\gamma^\mu\partial_\mu - M]\psi - g(\overline{\psi}\gamma^\mu\psi)\partial_\mu\varphi \tag{1.7.13}$$

and comes to the conclusion that Z_Λ in $\psi_{r,\Lambda}(x) = Z_\Lambda^{-1/2}\psi_\Lambda(x)$ *diverges* in the cutoff limit $\Lambda \to \infty$. However, in his treatment ([**Stro13**], Section 6.3), the wave-function renormalisation takes the form

$$Z_\Lambda(g) = e^{-i\frac{1}{2}g^2\Delta_\Lambda^+(0)}, \tag{1.7.14}$$

where he denotes by $i\Delta_\Lambda^+(x) = (2\pi)^{-3}\int_\Lambda d^4p\ \delta_+(p^2 - m^2)e^{-ip\cdot x}$ the two-point Wightman function[24] of the free field with a UV cutoff. Since this obviously implies

$$\lim_{\Lambda\to\infty} Z_\Lambda(g) = \lim_{\Lambda\to\infty} e^{-i\frac{1}{2}g^2\Delta_\Lambda^+(0)} = 0, \tag{1.7.15}$$

we do not exactly find that Z diverges, but he probably meant that is vanishes. However, this example is at least in accord with the divergence of the spectral integral (1.7.12) and can thus be reconciled with his claim $Z \to 0$ made for the scalar field.

To our mind, these discussions only point to the intricasies of the 'multiplicative renormalisation folklore': for the two-point function of the renormalised field, we find

$$\langle\Omega|\varphi_r(x)\varphi_r(y)\Omega\rangle = Z^{-1}\langle\Omega|\varphi(x)\varphi(y)\Omega\rangle \tag{1.7.16}$$

which is, for example, also used to derive the Callan-Symanzik equation[25]. Because the two-point function of the renormalised field φ_r is finite and that of the unrenormalised field φ diverges, Z needs to diverge and hence cannot vanish as proposed by Strocchi.

We opine that this contradiction shows *how poorly understood the connection between the Källen-Lehmann representation and canonical perturbation theory actually is*: in perturbation theory, Z is given as a perturbative series with respect to the renormalised coupling g_r (see Section 3.4) having two nasty properties:

- for finite regulator or cutoff, $Z(g_r)$ has an asymptotic power series (clearly divergent),
- what is more, its coefficients diverge when the regulator (or cutoff) is removed, exacerbating things for any attempt to understand it nonperturbatively through resummation schemes.

It can therefore never satisfy $0 \leq Z \leq 1$, which is a standard assertion in textbooks whenever the spectral representation (1.7.8) is derived. Take [**PeSch95**] for example. On page 215, they construe Z as "... the probability for $\phi(0)$ (that is, the field at $x = 0$, author's note) to create a given state from the vacuum." This is because in their analysis, they find

$$Z = |\langle\Omega|\phi(0)\lambda_0\rangle|^2 \tag{1.7.17}$$

[24] We introduce this function in Section 2.1.

[25] We will present a derivation of this equation in Chapter 5, Section 5.4 and show that within the Hopf-algebraic setting, there is a mathematically sounder yet more technical way to derive it.

where $|\lambda_0\rangle$ is the zero-momentum state of the interacting theory and, therefore, Z has to be within the unit interval. But their deliberations are typical and can in fact be traced back to [**BjoDre65**]. However, before elaborating on this, we express at this point our contention that if we were to conduct an anonymous survey amongst theoretical physicists on this issue, the outcome would certainly be a collection of contrived narratives[26].

1.7.4. Asymptotic scattering theory and wave-function renormalisation. In an attempt to connect two other disconnected stories of QFT, the renormalisation constant Z found its way also into *asymptotic scattering theory.* From studying the literature, we glean the following brief history of developments regarding this issue.

1. Lehmann, Symanzik and Zimmermann formulate a theory of scattering for quantised fields in [**LSZ55, LSZ57**] stating explicitly in their abstract that *"These equations contain no renormalization constants, but only experimental masses and coupling parameters. The main advantage over the conventional formalism is thus the elimination of all divergent terms in the basic equations. This means no renormalization problem arises."* And indeed, in these papers, no such constants appear: they introduce a spatially smeared scalar field (which they denote by $A(x)$, but never mind)

$$\varphi_f(t) := i \int d^3x \; [f^*(t,\mathbf{x})\partial_t\varphi(t,\mathbf{x}) - \partial_t f^*(t,\mathbf{x})\varphi(t,\mathbf{x})], \tag{1.7.18}$$

where $f(x)$ is a solution of the Klein-Gordon equation and integrable in some sense (see [**LSZ55**]). This is essentially supposed to form a 'quantised wave packet'. Then they state the *asymptotic condition* for the existence of a free incoming field as

$$\langle\alpha|\varphi_f(t)|\beta\rangle \sim \langle\alpha|\varphi_{f,in}(t)|\beta\rangle \qquad \text{as} \quad t \to -\infty \tag{1.7.19}$$

2. In contrast to this, however, Bjorken and Drell cite [**LSZ55**] in [**BjoDre65**] (Section 16.3) and rephrase the asymptotic condition in the form

$$\langle\alpha|\varphi_f(t)|\beta\rangle \sim Z^{1/2}\langle\alpha|\varphi_{f,in}(t)|\beta\rangle \qquad \text{as} \quad t \to -\infty \tag{1.7.20}$$

for the incoming field and for the outgoing accordingly. They have slipped in a 'normalisation factor' Z allegedly on the grounds that the matrix elements $\langle\alpha|\varphi_{in}(x)|\beta\rangle$ need to be 'normalised'. This is exactly the point where their exposition departs from [**LSZ55**], but to be fair, they take pedagogical care only to speak of a *normalisation* constant within the bounds of Chapter 16. But Bjorken and Drell's departure is completed also in spirit when they subject incoming Dirac fields to the same procedure in Section 16.8 and identify the corresponding normalisation constant with the wave-function *re*normalisation Z_2 in Section 19.7, where Chapter 19 is devoted to renormalisation.

They repeat at this point their interpretation of Z_2 as "the probability of finding a 'bare-electron' state within the one-electron state of the interacting theory". Here we see what great lengths physicists go to in order to make sense of everything.

As far as the CCR/CAR are concerned, they pervade this textbook: in Section 16.3 Bjorken and Drell demand that the interacting scalar field obey the CCR with $\pi(x) = \partial_t\varphi(x)$ and come in Section 16.4 on the spectral representation to the same conclusion[27] as Lopuszanski in (1.7.11).

---

[26] Nowadays, video lectures on QFT abound on the internet. For example, the Perimeter Institute based in Canada sports a vast library of such recordings. In one lecture on the spectral representation, given by a physicist whose name we will not give away, a student says "... but you get infinity!". After some silence lasting a lengthy 8 seconds (!), the lecturer resumes by speaking of UV divergences, and, gradually gaining back his normal speed of talking, he explains "... if you cut the thing off, it really is true (the statement $0 \leq Z \leq 1$, author's note), but if you push the cutoff to infinity, then it won't be true anymore."

[27] The reader be warned: the statement $Z^{-1} = 1 + \int d\mu^2\ \sigma(\mu^2)$ takes in [**BjoDre65**] a seemingly different form, namely $1 = Z + \int d\mu^2\ \sigma(\mu^2)$. The resolution is that their spectral function is unrenormalised, ie dividing by Z gives Lopuszanski's equation.

The reason why Bjorken and Drell placed the factor Z *before* the free incoming field $\varphi_{f,in}(t)$ is, in contrast to [**LSZ55**], that they let $\varphi_f(t)$ in (1.7.20) be the unrenormalised field such that for the renormalised field, one gets the asymptotic identity

$$\langle\alpha|\varphi_{r,f}(t)|\beta\rangle \sim Z^{-1/2}\langle\alpha|\varphi_f(t)|\beta\rangle \sim \langle\alpha|\varphi_{f,in}(t)|\beta\rangle \qquad \text{as} \quad t \to -\infty. \tag{1.7.21}$$

This is now the asymptotic condition of [**LSZ55**]. Not facilitating the comprehensibility of their arguments, however, they do not make this point clear anywhere in the text (which they could have in their chapter on renormalisation; great book though).

Scattering theory of Haag & Ruelle. However, the asymptotic condition of Lehmann, Symanzik and Zimmermann in [**LSZ55**] was of *axiomatic nature*. No proof of any kind that the fields obey this condition was given. Small wonder given that Wightman's axioms were not yet formulated at the time. This situation changed when, based on these axioms, the asymptotic condition was shown by Ruelle to be satisfied under some additional requirements [**Rue62**]. The proof had already essentially been worked out by Haag in [**Ha58**], albeit at a slightly lower level of rigour (according to Ruelle). This rigorous form of LSZ scattering theory was hence dubbed *Haag-Ruelle scattering theory* (see also Subsection 2.2.2).

1.8. What to do about Haag's theorem: reactions

After having discussed at length the CCR/CAR question for interacting fields, we come back to Haag's theorem. In perturbation theory, the wave-function renormalisation constant does its job impeccably. Because perturbative results are obtained without assuming the renormalised fields to conform with the CCR/CAR, perturbative QFT is unscathed by Powers' and Baumann's results.

In contrast to this situation, Haag's theorem stands in direct contradiction to the canonical formalism of perturbation theory which presupposes and relies on the existence of Dyson's matrix and the interaction picture.

Some say the price for ignoring this and other triviality theorems are the UV divergences that have beset the canonical formalism from the very start [**Wi67**]. While this is certainly true, one has to appreciate the role of *renormalisation* which has become an integral part of the formalism. This subtraction scheme does not only successfully fix these problems by rendering individual integrals finite and well-defined but changes the theory systematically.

One has to say that both Haag's and van Hove's papers [**Ha55, vHo52**] came out relatively late, in the early 1950s, that is, *after* the renormalisation problem of quantum electrodynamics (QED) had already been settled by Feynman, Tomonaga, Schwinger and Dyson (see [**Feyn49, Dys49a, Dys49b**] and references there). In hindsight it was perhaps fortunate that these pioneers did not get sidetracked by representation issues.

However, in Dyson's review of Haag's 1955 paper, he acknowledges that

> "The meaning of these results is to make even clearer than before the fact that the Hilbert space of ordinary quantum mechanics is too narrow a framework in which to give consistent definition to the operations of quantum field theory ... attempts to build a rigorous mathematical basis for field theory ... always stop short of any nontrivial examples. The question, what kind of enlarged framework would make consistent definitions possible, is the basic unsolved problem of the subject."

Aware of van Hove's work, he furthermore writes "The so-called 'Haag's theorem' ... is essentially an old theorem of L.van Hove"[28].

[28] Dyson's review of Haag's paper is available at www.ams.org/mathscinet, keywords: author "Haag", year 1955.

Clearly, trying to explain and tackle the puzzle posed by Haag's theorem certainly falls within the remit of *mathematical* physics which is probably why it was largely ignored by practising physicists or even dismissed. Källen is quoted saying that the result "... is really of a very trivial nature and it does not mean that the eigenvalues of a Hamiltonian never exist or anything that fundamental." [**Lu05**] Such attitudes inspired Wightman to write in a proceedings paper that "... there is a widespread opinion that the phenomena associated with Haag's theorem are somewhat pathological and irrelevant for real physics. I make one more attempt to explain why this is not the case." [**Wi67**]

1.8.1. Textbooks. Despite all this, nowadays' standard physics textbooks (eg [**PeSch95, Wein95**]) do not mention Haag's theorem in any way. To find a discussion of it, one has to turn to textbooks from before 1970 which treat it rather differently: while the books by Roman [**Ro69**] and Barton [**Bar63**] devote several pages to it, Bjorken and Drell [**BjoDre65**] say in a footnote on p.175 that although Haag's theorem excludes the existence of Dyson's matrix, they will (understandably) assume that it does exist notwithstanding.

Sterman takes a similiar stance in [**Ster93**] which he relegated to the appendix where he reviews the interaction picture. Her writes on p.508

> "Haag's theorem states that the unitary transformation ... is not strictly consistent with Poincarè invariance ... this fascinating point, ... however, has not been shown to affect practical results, ..."

This is probably the most down-to-earth standpoint possible. Ticciati's view in [**Ti99**], p.84 is somewhat more subtle,

> "[the] first few terms yield wonderfully accurate predictions. It appears then that the interaction picture provides a sound approach to perturbation theory but may have no non-perturbative validity."

This gives the unfortunate impression that the author does not take the results of renormalised perturbation theory seriously. However, Ticciati suggests that one may drop the assumption of unitary equivalence and, also on the plus side, his formulation of Haag's theorem corresponds to our version in that he does not make use of the conjugate momentum fields.

Barton. Barton discusses the technicalities of the Wightman distributions' analyticity properties and surveys the provisions of Haag's theorem at considerable length to fathom out which may be relinquished. The one he would like to retain is the condition that both fields obey the CCR (see [**Bar63**], p.158). He includes this property in his exposition of Haag's theorem as part of the conditions for the generalised theorem (1.2.18) (ibid., p.153), exactly where it has no business to hang around: not surprising therefore that he neither uses the CCR nor gives reasons why they are needed. In fact, the reason he does not employ them is that they are dispensible (see [**StreatWi00**], p.100) and, as we have seen in the preceding section, probably not fulfilled by a theory of true interactions. To put this in perspective, however, Barton's book [**Bar63**] clearly shows that at the time of writing, its author favoured the Yang-Feldman approach to the S-matrix in the Heisenberg picture which we mentioned in the context of Lopuszanski's work in the previous section in which the CCR are an integral part (Yang and Feldman's seminal paper is [**YaFe50**]).

Roman. While Roman's account of Haag's theorem and his observation that the CCR may be discarded are entirely correct ([**Ro69**], pp.330 and pp.392), we do not at all agree with his assertions as to what stand to take on this issue [**Ro69**]: the author contends that *renormalisation does not help.*

We have to mention that the author seemed to have a weird understanding of the Stone-von Neumann theorem (Theorem 1.1 of this thesis) and deemed it applicable to QFT ([**Ro69**], p.330). According to his understanding, it is merely the assertion that unitary equivalence of two field algebras implies that both algebras are irreducible, if one is. While this assertion is hardly questionable, we know that the real Stone-von Neumann theorem does not hold for an infinite number of degrees of freedom and hence not for QFT (see Sections 1.1,1.2). However, to make his case about renormalisation, he (wrongly) invokes his understanding of the theorem to argue as follows.

Assuming that by the Stone-von Neumann theorem the CCR only permit unitarily equivalent representations, the wavefunction (=field-strength) renormalisation Z in the CCR

$$[\varphi(\mathbf{x}), \pi(\mathbf{y})] = iZ^{-1}\delta^{(3)}(\mathbf{x}-\mathbf{y}) \tag{1.8.1}$$

of the renormalised field $\varphi(\mathbf{x})$ does not change the fact that this 'interacting' renormalised field is free. Apart from the fact that the wavefunction renormalisation constant Z is a dubious fellow which forces in our opinion the rhs of (1.8.1) to vanish, he ignores that there is much more intricate formalism between the free and the interacting field than the canonical formalism pretends.

This is one central assertion of this thesis which we shall discuss in Sections 3.4 and 3.5: although the canonical apparatus speaks of a unitary map between the free interaction picture and the interacting Heisenberg picture field, it disproves what it purports by its own actions. Encountering divergences, there is a subsequent backpedalling and fiddling in of infinite factors, only to claim again unitarity equivalence where there cannot be any (we will next-to prove it within the mindset of the canonical formalism in Section 3.5).

However, on the plus side, Roman discusses the provisions of Haag's theorem and says that he is inclined to give up unitary equivalence and maybe also the CCR (ibidem, p.393). Apart from his peculiar idea of what the Stone-von Neumann theorem is about, he makes valuable remarks on the CCR and its relation to the property of locality (ibidem, p.328): if locality holds, ie

$$[\varphi(x), \varphi(y)] = 0 \qquad (x-y)^2 < 0, \tag{1.8.2}$$

then the CCR must be singular because

$$\frac{1}{\varepsilon}[\varphi(t,\mathbf{x}), \varphi(t+\varepsilon,\mathbf{y}) - \varphi(t,\mathbf{y})] = 0, \tag{1.8.3}$$

for all small enough $\varepsilon \neq 0$. He also mentions that we cannot be sure that this singular commutator is a c-number. Despite his interest in the CCR, Roman was not aware of Powers' results, which were published in 1967, two years prior to the publication of his book. He was probably spared this news[29].

1.8.2. Monographs. We have found two monographs covering Haag's theorem. One of them, [**Stro13**] by Strocchi, has already been cited numerous times in this work. Let us first consider the other one.

Duncan. The notable recent extensive and mathematically-oriented monograph by Duncan, [**Dunc21**], contains a section on Haag's theorem entitled "How to stop worrying about Haag's theorem". The author shows that a simple mass shift leads to the van Hove phenomenon, ie the vanishing overlap of the two vacua.

This can actually be tracked down to the fact that free field theories with different masses are unitarily inequivalent, which he seemed not to be aware of when he wrote his book (see Theorem 3.1 or Theorem X.46 in [**ReSi75**], p.233). Most interestingly, he subsequently treats the mass shift as a 'mass perturbation', ie the difference between the two masses as an interaction term

[29]His discussion of the CCR would have been much more extensive; historical remark: a case of disconnected research communities although Roman, as far as one can tell from his textbook was clearly mathematically inclined.

and applies standard perturbation theory in the interaction picture to it. The upshot there is that the propagator of the mass-shifted field comes out correctly.

This brings us right to the very contention underlying this thesis on which we shall elaborate in Section 3.5: *unitary inequivalence does not mean that canonical (renormalised) perturbation theory yields nonsensical and unphysical results*. In fact, we use Duncan's example to conclude that *renormalisation allows us to evade Haag's triviality dictum*!

Duncan points out that Haag's theorem does not exclude the existence of the S-matrix maintaining that both Haag-Ruelle and LSZ scattering theories "lead to a perfectly well-defined, and unitary, S-matrix" on the basis of the axiomatic framework ([**Dunc21**], p.364). The author then explains his attitude towards this issue ([**Dunc21**], pp.369-370):

- one should in a FIRST STEP introduce a spatial volume and a UV cutoff to work with a finite number of degrees of freedom and hence a well-defined interaction picture,

while accepting the price of sacrificing Poincarè invariance at this stage. Then, having kept Haag's theorem at bay so far by retaining the provisions of the Stone-von Neumann theorem,

- one redefines in a SECOND STEP both mass and couplings by renormalisation and removes all cutoffs which finally restores Poincarè invariance.

The upshot is, at least according to Duncan, that we have enough reason to stop worrying about Haag's theorem because we have by virtue of this procedure circumvented it. This is essentially the view that nowadays we believe practising physicists generally subscribe to or would subscribe to if we had told them this story.

Yet in the ambition to retain the interaction picture as long as possible along the way, however, the author leaves the impression that he wishes to arrive at a unitarily equivalent representation of the CCR. As we have seen in the previous section and will extensively discuss in Section 3.4, this is certainly not where the 'circumventing' procedure leads. If the aim of a 'circumvention scheme' is unitary equivalence to free fields, then it has to go wrong as Haag's theorem cannot be circumvented in this sense: it is a mathematical theorem in the truest sense of the word; it brings with it the 'hardness of the logical must'[30].

Strocchi. Strocchi puts the triviality dictum expressed through Haag's theorem down to the facts we have discussed and explained in Section 1.5: the fact that any theory within the unitary equivalence class of the free particle Fock space $\mathfrak{H}_0$ must have a unitarily equivalent generator of time translations H_0 is irreconcilable with a splitting into a free and another nontrivial piece.

Strocchi sees the interaction picture as instrumental in computing nontrivial results in perturbation theory but says unitary equivalence to the Heisenberg field ceases to make sense in the no-cutoff limit, even after renormalisation ([**Stro13**], pp.52). We agree with this fully and have to stress again that Haag's theorem cannot be circumvented in this way: Haag's theorem is a general statement and does not take into account any feature of Dyson's matrix, ie the field intertwiner, other than its unitarity. All other provisions have a fundamentally different status, as we shall discuss in Section 2.3 where we present a proof of Haag's theorem.

1.8.3. Papers. Guenin and Segal suggest that Haag's theorem can be bypassed if the time evolution of the interaction picture is implemented in the form of *locally* unitary automorphisms and not, as usual, by globally unitary maps [**Gue66, Seg67**].

The former author introduces a modified interaction picture in which the trivial part of the Hamiltonian acts on the states and the nontrivial one on the observables. The Dyson series he then obtains for a specific class of Euclidian invariant Hamiltonians does, however, to say the least, not only look peculiar but also leads to a convergent perturbation series in one space dimension[31].

[30] Wittgenstein

[31] Recall that $(\varphi^4)_2$ has an asymptotic perturbation series and, as is well-known, so does $(\varphi^4)_0$! [**GliJaf81**].

1.8.4. Fraser's thesis. The most extensive review of Haag's theorem is Doreen Fraser's doctoral thesis [**Fra06**]. She belongs to a small community of philosophers of physics whose work revolves around the interpretation of quantum theory in general and, in her case, of quantum field theory in particular. Her thesis is roughly composed of 3 parts.

In the first, she expounds Haag's theorem and sketches its proof. In the second part, she discusses possible responses, eg

- introduction of a volume cutoff at the price of sacrificing translational invariance,
- renormalisation, discussed using the example of $(\varphi^4)_2$, characterised by the cutoff limit

$$H_{\mathrm{ren}} = \lim_{\Lambda\to\infty} \{H_\Lambda - E^0_\Lambda\}, \tag{1.8.4}$$

  which comes at the price of an ill-defined counterterm for the ground state energy E^0_∞,
- dropping the assumption of unitary equivalence and using other approaches like Haag-Ruelle scattering theory and constructive approaches.

Our view on these points is that none of them is satisfactory: while the constructive method is fine unfortunately only for lower-dimensional Minkowski spacetime with more or less trivial rotations (see also Subsection 1.4.2), Haag-Ruelle theory is impractical: no predictions can be made and the connection to renormalised QFT is still unclear (see Subsection 1.7.4).

The third part is devoted to ontological questions. Her conclusion about the ontological status of particlelike entities is based on the tenuous mode of existence of particles as described by interacting non-Fock QFTs:

> " ... since in the real world there are always interactions, QFT does not furnish grounds for regarding particlelike entities as fundamental constituents of reality"([**Fra06**], p.137).

This is, from a physics point of view, too radical since *elastic interactions* are also captured by QFT, albeit in a nonrelativistic limit. However, here is an interesting aspect: Fraser states that Haag's theorem is not associated with UV but infinite-volume divergences!

In a way, this strikes us as plausible: Haag's theorem is FIRSTLY independent of the dimension of spacetime[32], which is why superrenormalisable theories are also affected by Haag's theorem and SECONDLY Wightman's arguments described in Subsection 1.5.2 against the interaction picture are clearly based on an infinite-volume divergence.

Her statement that "...Haag's theorem undercuts global unitary equivalence, it is compatible with local unitarity equivalence."([**EaFra06**], p.323) supports this view, which is inspired by a modified version of Theorem 3.1 ('Haag's theorem for free fields', Subsection 3.5.1) with a contrary outcome ([**ReSi75**], p.329): if φ_0 and φ are two free fields with masses m_0 and m, respectively, and $B \subset \mathbb{M}$ a bounded region, then there exists a unitary map $V_B\colon \mathfrak{H}_0 \to \mathfrak{H}$ between the two corresponding Hilbert spaces such that $\varphi(f) = V_B\varphi_0(f)V_B^{-1}$ for all $f \in \mathscr{D}(B)$. This fits in nicely with Guenin and Segal's studies (see above, Subsection 1.8.3).

But when proper interactions enter the game in physical Minkowski space $\mathbb{M} = \mathbb{R}^4$, one cannot separate the UV divergences away and declare them as having nothing to do with Haag's theorem. We prefer a more modest position regarding this issue: the ill-definedness of the interacting picture in any spacetime leads to several types of divergences.

HAAG'S THEOREM SPEAKS OF NO DIVERGENCES. IT HAS TO BE MERELY READ AS SAYING: EITHER BOTH THEORIES ARE FREE OR THEY MUST BE UNITARILY INEQUIVALENT.

[32] It should be a QFT, ie spacetime should be at least two-dimensional! We want Lorentz boosts!

CHAPTER 2

Axiomatics and proof of Haag's theorem

As is well-known, the canonical formalism is no mathematically coherent framework. For, if one tries to translate its notions, in particular the idea of an *operator field*, into the language of operator theory, mathematical inconsistencies arise. The general pattern of the associated problems is that whenever an object is overfraught with conditions, a canonical computation brings about nonsensical results, while a strict mathematical treatment finds that the object must be trivial to maintain well-definedness.

We shall see in Section 2.1 how this manifests itself already in the case of a free scalar field. A triviality result by Wightman presented there is of interest because it prepares the ground for the axioms discussed in Section 2.2. These axioms, known as *Wightman axioms*, comprise the framework on which the proof of Haag's theorem relies. As we go along, we will mention their tenuous status when it comes to circumscribing what QFT should be about.

Next, we present a bunch of pertinent results on scalar fields that adhere to these axioms in the first part of Section 2.3 to set the stage for the proof of Haag's theorem in the second part. Describing its provisions in detail, we will see that unitary equivalence is one of the core conditions.

Because Haag's theorem is in the literature generally only formulated for scalar fields, we investigate the case of fermion and gauge fields in Section 2.4. While it goes through trivially for fermions, gauge theories raise hard questions that originate in a fundamental incompatibility of the axiomatic framework with quantum electrodynamics (QED). The pertinent issues are detailed and a conclusion about whether Haag's theorem applies to QED is drawn.

2.1. Canonical quantum fields: too singular to be nontrivial

Let us start with an innocent-looking canonical free Hermitian scalar field $\varphi(x)$, given formally by its Fourier expansion

$$\varphi(x) = \int \frac{d^4p}{(2\pi)^3} \frac{1}{\sqrt{2E_p}} \, [e^{-ip\cdot x} a(\mathbf{p}) + e^{ip\cdot x} a^\dagger(\mathbf{p})], \tag{2.1.1}$$

where $p_0 = E_p = \sqrt{\mathbf{p}^2 + m^2}$ is the energy of the scalar particle. The mode opersators satisfy $a(\mathbf{p})\Psi_0 = 0$ and

$$[a(\mathbf{p}), a(\mathbf{q})] = 0 = [a^\dagger(\mathbf{p}), a^\dagger(\mathbf{q})] \ , \qquad [a(\mathbf{p}), a^\dagger(\mathbf{q})] = i\delta^{(3)}(\mathbf{p} - \mathbf{q}), \tag{2.1.2}$$

as usual. The trouble starts as soon as we ask for the norm of the 'state' $\Psi = \varphi(x)\Psi_0$. If the canonical field $\varphi(x)$ is to be taken seriously as an operator at a sharp spacetime point $x \in \mathbb{M}$, this should be a valid question. However, applying (2.1.2), we find $||\Psi|| = ||\varphi(x)\Psi_0|| = \infty$. The expedient is to smooth out the field with respect to its spatial coordinates, as in

$$\varphi(t, f) = \int d^3x \ f(\mathbf{x})\varphi(t, \mathbf{x}), \tag{2.1.3}$$

where $f \in \mathscr{S}(\mathbb{R}^3)$ is a Schwartz function. If we now compute the norm of the state vector $\Psi_f(t) = \varphi(t,f)\Psi_0$ we find

$$||\Psi_f(t)||^2 = ||\varphi(t,f)\Psi_0||^2 = \langle \Psi|\varphi(t,f^*)\varphi(t,f)\Psi_0\rangle = \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\frac{|\widetilde{f}(\mathbf{p})|^2}{\sqrt{\mathbf{p}^2+m^2}}. \tag{2.1.4}$$

This integral is convergent on account of $\widetilde{f}$ being Schwartz, where $\widetilde{f} = \mathscr{F}f$ is the Fourier transform of the Schwartz function[1] $f \in \mathscr{S}(\mathbb{R}^3)$. Consequently, the state $\Psi_f(t) = \varphi(t,f)\Psi_0$ has finite norm and exists. But this does not hold for the state $\varphi(x)\Psi_0$.

We conclude that the canonical free field needs smearing at least in space to be well-defined on the vacuum. Without smearing, it is merely a symbol. Nonetheless, computing the two-point function

$$\langle \Psi_0|\varphi(x)\varphi(y)\Psi_0\rangle = \frac{1}{2}\int \frac{d^3p}{(2\pi)^3}\frac{e^{-ip\cdot(x-y)}}{\sqrt{\mathbf{p}^2+m^2}} =: \Delta_+(x-y;m^2) \tag{2.1.5}$$

yields a well-defined function for $x \neq y$, but has a pole where $x = y$, ie a short-distance singularity.

2.1.1. Triviality of sharp-spacetime fields. The following pertinent theorem due to Wightman says that if one assumes a quantum field $\varphi(x)$ exists as an operator at a sharp spacetime point $x \in \mathbb{M}$ and is covariant with respect to a strongly continuous representation of the Poincarè group, then it is trivial in the sense that it is merely a multiple of the identity [**Wi64**]:

THEOREM 2.1 (Short distance singularities). *Let $\varphi(x)$ be a Poincaré-covariant Hermitian scalar field, that is,*

$$U(a,\Lambda)\varphi(x)U(a,\Lambda)^\dagger = \varphi(\Lambda x + a) \tag{2.1.6}$$

and suppose it is a well-defined operator with the vacuum Ψ_0 in its domain. Then the function

$$F(x,y) = \langle \Psi_0|\varphi(x)\varphi(y)\Psi_0\rangle \tag{2.1.7}$$

is constant, call it c. Furthermore $\varphi(x)\Psi_0 = \sqrt{c}\Psi_0$, ie $\varphi(x)$ is trivial and thus

$$\langle \Psi_0|\varphi(x_1)...\varphi(x_n)\Psi_0\rangle = c^{n/2}. \tag{2.1.8}$$

PROOF. We follow [**Stro13**]. First note that Poincaré covariance (2.1.6) implies

$$F(x+a, y+a) = F(x,y) \tag{2.1.9}$$

which entails that this function depends only on $(x-y)$. We write $F(x,y) = F(x-y)$. $F(x)$ is continuous by the strong continuity of the Poincarè representation, ie through the covariance identity $\varphi(x) = U(x,1)\varphi(0)U(x,1)^\dagger$ inserted into the two-point function,

$$F(x) = \langle \Psi_0|\varphi(x)\varphi(0)\Psi_0\rangle = \langle \varphi(0)\Psi_0|U(x,1)^\dagger\varphi(0)\Psi_0\rangle = \langle \varphi(0)\Psi_0|U(x,1)\varphi(0)\Psi_0\rangle^* \ . \tag{2.1.10}$$

By virtue of this and the property $\int d^4x \int d^4y \ f(x)^* F(x-y) f(y) = ||\varphi(f)\Psi_0||^2 \geq 0$ with completely smoothed-out field $\varphi(f) := \int d^4x f(x)\varphi(x)$, we conclude that $F(x)$ is a continuous function of positive type. The Bochner-Schwartz theorem tells us now that there exists a positive tempered measure μ on $\mathbb{R}^4$ such that

$$F(x) = \int e^{-ip\cdot x} \ d\mu(p), \tag{2.1.11}$$

[1] An explicit and careful treatment starts with $f \in \mathscr{D}(\mathbb{R}^3)$ to ensure the theorem of Fubini can be employed and then extends the result to $\mathscr{S}(\mathbb{R}^3)$.

ie $F(x)$ is the Fourier transform of a positive tempered measure. If we use the spectral representation of the translation symmetry operator,

$$U(x,1) = \int e^{ip\cdot x} d\mathsf{E}(p), \tag{2.1.12}$$

plug it into (2.1.10), we see that the measure μ must be Poincaré invariant by $F(\Lambda x) = F(x)$ and

$$F(x) = \langle \Psi_0 | \varphi(x)\varphi(0)\Psi_0\rangle = \int e^{-ip\cdot x} d\langle \varphi(0)\Psi_0|\mathsf{E}(p)\varphi(0)\Psi_0\rangle. \tag{2.1.13}$$

Then it follows that μ is of the form ([**ReSi75**], p.70)

$$d\mu(p) = c\ \delta^{(4)}(p) d^4p + b\ dm^2\ d^3p\ \frac{\rho(m^2)}{\sqrt{\mathbf{p}^2+m^2}} \qquad (c, b \geq 0), \tag{2.1.14}$$

which is essentially the Källen-Lehmann spectral representation with spectral function

$$\rho(m^2) \geq 0\ , \qquad \operatorname{supp}(\rho) \subset [0,\infty). \tag{2.1.15}$$

Because $F(x)$ is continuous at $x = 0$, we have

$$F(0) = \int d\mu(p) = c + b \int dm^2 \int d^3p\ \frac{\rho(m^2)}{\sqrt{\mathbf{p}^2+m^2}} \tag{2.1.16}$$

which implies $b = 0$ because of the UV divergence of the momentum integral. This means in particular $F(x) = F(0) = c$ and that the spectral measure $d\mathsf{E}(p)$ has support only at $p = 0$, where $\mathsf{E}(0) = \langle \Psi_0 |\ \cdot\ \rangle \Psi_0$, ie where it projects onto the vacuum (the vacuum is the only state with vanishing energy). If we write this in terms of (2.1.13), we get

$$c = F(0) = \int d\langle \varphi(0)\Psi_0|\mathsf{E}(p)\varphi(0)\Psi_0\rangle = \langle \varphi(0)\Psi_0|\Psi_0\rangle\langle\Psi_0|\varphi(0)\Psi_0\rangle = |\langle\Psi_0|\varphi(0)\Psi_0\rangle|^2. \tag{2.1.17}$$

Using this, one easily computes $||(\varphi(x) - \sqrt{c})\Psi_0||^2 = 0$. □

This result is insightful. We know exactly which assumption cannot be true for the two-point function of the canonical free field in (2.1.1) and are even able to put our finger on it: the function $F(x-y) = \Delta_+(x-y; m^2)$ in (2.1.5) is not continuous at $x - y = 0$ as the integral diverges logarithmically in this case. The argument that led to this assumption can be easily traced back to the condition of strong continuity of the representation of the translation group, ie the requirement that the function

$$x \mapsto \langle \Psi | U(x,1)\Phi\rangle = \langle \Phi | U(x,1)^\dagger \Psi\rangle^* \tag{2.1.18}$$

be a continuous function for all state vectors $\Psi, \Phi \in \mathfrak{H}$. The erroneous assumption for our free field is therefore (as we know) that the state $\Psi = \Phi = \varphi(0)\Psi_0$ is one of these permissible state vectors.

Interestingly enough, there is an analogy to Haag's theorem.

- First of all the assumption that there exists a Poincarè-covariant sharp-spacetime field is too strong.
- Secondly, while the rigorous procedure takes well-reasoned steps and ends up with pleading triviality, the formal canonical calculation leads to an infinite result.

This is also exactly what happens in non-renormalised canonical perturbation theory which, by its very nature, has to work with sharp-spacetime fields.

2.1.2. Tempered distributions. However, Wightman's theorem is not applicable to the smoothed-out free field $\varphi(t,f)$ in (2.1.3). One reason is that Poincarè covariance cannot be formulated like in (2.1.6) but has to be altered, in particular, time must also be smeared.

The axiomatic approach to be introduced in the next section proposes to construe the two-point function (2.1.5) as a symbol for a tempered distribution, ie

$$\mathscr{S}(\mathbb{M}) \times \mathscr{S}(\mathbb{M}) \ni (f,g) \mapsto W(f,g) = \int d^4x \int d^4y \; f^*(x)W(x-y)g(y). \tag{2.1.19}$$

This amounts to defining a 'two-point' distribution $W \in (\mathscr{S}(\mathbb{M}) \times \mathscr{S}(\mathbb{M}))'$ by

$$W(f,g) := \langle \Psi_0 | \varphi(f)\varphi(g)\Psi_0\rangle \tag{2.1.20}$$

with completely smooth-out field opersators $\varphi(f) = \int d^4x \; f(x)\varphi(x)$ and $\varphi(g) = \int d^4x \; g(x)\varphi(x)$. Because $\varphi(f)$ makes sense as an operator and gives rise to distributions, Wightman called these objects *operator-valued distributions.*

Note that from a conceptual and physical point of view, the smoothing operation imposes no restriction. Observable fields cannot be measured with arbitary precision, and smearing a field in both time and space with a test function of arbitarily small support is certainly permissible and not a big ask. This had actually been realised already much earlier by Bohr and Rosenfeld [**BoRo33, BoRo50**].

The smoothing has the nice effect that (2.1.19) can be written in Fourier space as

$$W(f,g) = \int \frac{d^4p}{(2\pi)^4} \; \widetilde{f}^*(p)\widetilde{W}(p)\widetilde{g}(p) =: \widetilde{W}(\widetilde{f},\widetilde{g}) \tag{2.1.21}$$

and that the too strong assumption of continuity of $F(x) = W(x)$ at $x = 0$ in (2.1.16) can now be replaced by the much weaker condition that $d\mu(p) = \widetilde{W}(p)d^4p$ be a well-defined Poincarè invariant distribution, ie

$$\int h(p) \; d\mu(p) = c \; h(0) + b \int d\nu^2 \int d^3p \; \frac{\rho(\nu^2)}{\sqrt{\mathbf{p}^2 + \nu^2}} \; h(p) \tag{2.1.22}$$

for a Schwartz function $h \in \mathscr{S}(\mathbb{M})$. Then, with this weaker requirement, the integral on the rhs of (2.1.22) need not be muted, ie the choice $b \neq 0$ is perfectly acceptable unless the spectral function $\rho(\nu^2)$ goes berserk and overpowers the factor $1/|\nu|$.

The spectral representation of the free field's two-point distribution can be gleaned from comparing (2.1.5) with (2.1.22): we read off $c = 0$, $b = 1/(2(2\pi)^3)$ and $\rho(\nu^2) = \delta(\nu^2 - m^2)$.

Taking into account the singular nature of sharp-spacetime fields, Poincaré covariance (2.1.6) is reformulated for the smeared fields as

$$U(a,\Lambda)\varphi(f)U(a,\Lambda)^\dagger = \varphi(\{a,\Lambda\}f), \tag{2.1.23}$$

where $(\{a,\Lambda\}f)(x) = f(\Lambda^{-1}(x-a))$ is the transformed Schwartz function. As the reader may remember from Chapter 1.2, Haag's theorem relies on the sharp-spacetime version (2.1.6) of Poincarè covariance and cannot be applied to smeared fields. The reason is that the time of unitary equivalence is fixed and sharp, not averaged.

2.1.3. Sharp-time fields. We know from (2.1.4) that a free scalar field need only be smeared with respect to space to become a well-defined object. No one can tell whether this is actually the case for general (interacting) fields and some doubt it, eg Glimm and Jaffe ([**GliJaf70**], p.380) and Wightman ([**StreatWi00**], p.101). The latter authors speaks of 'examples' which suggest this but does not give a reference for further reading.

Powers and Baumann make use of 'relativistic' sharp-time fields in [**Pow67, Bau87, Bau88**] in the following sense. Starting with an operator-valued distribution transforming under the

Poincaré group as in (2.1.23), Baummann demanded that for a Dirac sequence $\delta_t^\epsilon \in \mathscr{S}(\mathbb{R})$ centred at time $t \in \mathbb{R}$ and any Schwartz function $f \in \mathscr{S}(\mathbb{R}^n)$ in space dimension n, the limit

$$\varphi(t, f) := \lim_{\epsilon \to 0} \varphi(\delta_t^\epsilon \otimes f) \tag{2.1.24}$$

exist, where $\mathbb{M} = \mathbb{R}^{n+1}$.

2.2. Wightman axioms and reconstruction theorem

Considering the issues incurred by working with sharp-spacetime and possibly also with sharp-time fields, it is no wonder that the following axioms due to Wightman and collaborators do not demand that general (interacting) quantum fields make sense as Hilbert space operators at sharp-spacetime points $x = (t, \mathbf{x})$ but only as operator-valued distributions.

2.2.1. Axioms for operator-valued distributions. Many authors quote [**Wi56**] as a seminal paper for the Wightman axioms. This is strictly speaking not true, as Wightman did not state them as such in this publication. He rather juggled with a few features that a reasonable QFT should bear without so easily falling prey to triviality results like the one illustrated by Theorem 2.1 in the previous section.

In [**Wi56**], Wightman first investigates the consequences that relativistic covariance, local commutativity and positivity of the generator of time translations entail for the vacuum expectation values of scalar fields. He then discusses how these properties suffice to 'reconstruct' the theory (reconstruction theorem). However, he tentatively adds that a 'completeness requirement' should be fulfilled to recover the entire theory. This requirement is now part of the axioms as *cyclicity of the vacuum* to be explained in due course.

The axioms were first explicitly enunciated by Wightman and Gårding in an extensive article [**WiGa64**] where they report on their reluctance to publish their results earlier. Although believing in their axioms' worth, they first wanted to make sure that nontrivial examples including free fields exist.

Except for the numbering, we follow [**StreatWi00**] in their exposition of the Wightman axioms. Although formulated for general quantum fields with any spin in their monograph, one has to say that *the axioms can only be expected to hold for scalar and Dirac fields.* For photon fields, the axiom of Poincaré invariance turned out to be incompatible with the equations of motion for free photons, ie Maxwell's equations for the vacuum. This is, of course, not the case for classical photon fields [**Stro13**]. So in hindsight, it was certainly a bit premature to include vector fields.

We shall describe the issues arising for gauge theories in Section 2.4 and in this section content ourselves with brief remarks. However, Wightman's axioms do not speak of any equation of motion for the fields to satisfy. From this perspective, issues arising with Maxwell's equations can be ignored. Since conventional quantisation schemes for free fields always involve equations of motion, one is reluctant to assent to this.

But because free scalar and a vast class of superrenormalisable QFTs conform with these axioms [**GliJaf81**], we expect them to make sense at least for scalar and fermion fields.

The axioms are organised in such a way that only the first one stands independently whereas the others that follow rely increasingly on the ones stated before. Here they are.

- **Axiom O (Relativistic Hilbert space)**. The states of the physical system are described by (unit rays of) vectors in a separable Hilbert space $\mathfrak{H}$ equipped with a strongly continuous unitary representation $(a, \Lambda) \mapsto U(a, \Lambda)$ of the connected Poincaré group $\mathscr{P}_+^\uparrow$. Moreover, there is a unique state $\Psi_0 \in \mathfrak{H}$, called the *vacuum*, which is invariant under this representation, ie

$$U(a, \Lambda)\Psi_0 = \Psi_0 \qquad \text{for all } (a, \Lambda) \in \mathscr{P}_+^\uparrow. \tag{2.2.1}$$

This first axiom merely sets the stage for a relativistic quantum theory without specifying any operators other than those needed for the representation of the Poincaré group. It therefore has the number O. For photons, however, this is already problematic: the Hilbert space must be replaced by a complex vector space with a nondegenerate inner product which is a much weaker requirement (see Section 2.4, or [**Stei00**], for example). The next axiom ensures that the Lorentz group cannot create an unphysical state, eg by sending a particle on a journey back in time.

- **Axiom I (Spectral condition)** The generator of the translation subgroup

$$i\frac{\partial}{\partial a^\mu}U(a,1)|_{a=0} = P_\mu \tag{2.2.2}$$

  has its spectrum inside the closed forward light cone: $\sigma(P) \subset \overline{V}_+$ and $H = P_0 \geq 0$, ie the time translation generator (=Hamiltonian) has nonnegative eigenvalues.

This axiom includes massless fields, ie fields without a mass gap. While Axiom I seems fine at face value, it is in fact at odds with QED and raises serious questions for a general canonical QFT on account of the consequences it has in store for the vacuum expectation values. We shall come back to this point below. The rest of the axioms introduce the concept of quantum fields and what properties they should have.

- **Axiom II (Quantum fields)**. For every Schwartz function $f \in \mathscr{S}(\mathbb{M})$ there are operators $\varphi_1(f), ..., \varphi_n(f)$ and their adjoints $\varphi_1(f)^\dagger, ..., \varphi_n(f)^\dagger$ on $\mathfrak{H}$ such that the polynomial algebra

$$\mathcal{A}(\mathbb{M}) = \langle\, \varphi_j(f), \varphi_j(f)^\dagger : f \in \mathscr{S}(\mathbb{M}), j = 1, ..., n \,\rangle_{\mathbb{C}} \tag{2.2.3}$$

  has a stable common dense domain $\mathfrak{D} \subset \mathfrak{H}$, ie $\mathcal{A}(\mathbb{M})\mathfrak{D} \subset \mathfrak{D}$ which is also Poincaré-stable, ie $U(\mathscr{P}_+^\uparrow)\mathfrak{D} \subset \mathfrak{D}$. The assignment $f \mapsto \varphi_j(f)$ is called *quantum field.* Additionally, the vacuum Ψ_0 is *cyclic* for $\mathcal{A}(\mathbb{M})$ with respect to $\mathfrak{H}$. This means $\Psi_0 \in \mathfrak{D}$ and the subspace

$$\mathfrak{D}_0 := \mathcal{A}(\mathbb{M})\Psi_0 \subseteq \mathfrak{D} \tag{2.2.4}$$

  is dense in $\mathfrak{H}$. Furthermore, the maps

$$f \mapsto \langle \Psi | \varphi_j(f) \Psi' \rangle \qquad (j = 1, ..., n) \tag{2.2.5}$$

  are tempered distributions on $\mathscr{S}(\mathbb{M})$ for all $\Psi, \Psi' \in \mathfrak{D}$.

As already alluded to in the previous section, it is due to this latter property, that a quantum field is referred to as an *operator-valued distribution.* The canonical notion of a quantum field can be approximated by the assignment of a spacetime point $x \in \mathbb{M}$ to an operator $\varphi_j(f_x)$ with a Schwartz function f_x of compact support in a tiny (Euklidean) ε-ball around $x \in \mathbb{M}$. This avoids the aforementioned ills of sharp-spacetime fields. Note that the operators in $\mathcal{A}(\mathbb{M})$ are not required to be bounded.

Cyclicity (2.2.4) expresses the condition that every (physical) state can be approximated to an arbitrarily high degree by applying the field variables to the vacuum. On account of the density of the so-obtained subspace $\mathfrak{D}_0$, one can then, if necessary, reach any state in $\mathfrak{H}$ after completion of $\mathfrak{D}_0$ with respect to Cauchy sequences. Finally, nullifying all zero norm states eliminates unphysical remnants.

It is important to note that this property, together with the following remaining axioms, is equivalent to *irreducibility* of the operator algebra $\mathcal{A}(\mathbb{M})$. This means that if C is an operator commuting with all field operators, then it is trivial, ie

$$[\mathcal{A}(\mathbb{M}), C] = 0 \qquad \Rightarrow \qquad C = c_0 1, \tag{2.2.6}$$

where $c_0 \in \mathbb{C}$. The proof can be found in [**StreatWi00**], Theorem 4-5.

- **Axiom III (Poincaré covariance)** The quantum fields transform under the (unitary representation of the) Poincaré group according to

$$U(a,\Lambda)\varphi_j(f)U(a,\Lambda)^\dagger = \sum_{l=1}^{n} S_{jl}(\Lambda^{-1})\varphi_l(\{a,\Lambda\}f), \tag{2.2.7}$$

on the domain $\mathfrak{D}$ where $S(\Lambda^{-1})$ is a finite-dimensional representation of the connected Lorentz group $\mathscr{L}_+^\uparrow$ and

$$(\{a,\Lambda\}f)(x) := f(\Lambda^{-1}(x-a)) \tag{2.2.8}$$

is the Poincaré-transformed test function.

For scalar fields, this takes the simple form $S(\Lambda^{-1}) = 1$, ie $S_{jl}(\Lambda^{-1}) = \delta_{jl}$, as we have seen in (2.1.23). The next property is called *locality* among proponents of the axiomatic approach and mostly *(Einstein) causality* or *microcausality* by practising physicists.

- **Axiom IV (Locality, Causality)**. Let $f, g \in \mathscr{S}(\mathbb{M})$ be of mutually spacelike-separated support, ie $f(x)g(y) \neq 0$ implies $(x-y)^2 < 0$. Then,

$$[\varphi_j(f), \varphi_l(g)]_\pm = \varphi_j(f)\varphi_l(g) \pm \varphi_l(g)\varphi_j(f) = 0, \tag{2.2.9}$$

for all indices (anticommutator '$+$' for fermions and commutator '$-$' for bosons).

This last axiom accounts for the fact that signals cannot travel faster than light, ie measurements at two different points in spacetime with spacelike separation do not interfere. With this interpretation, however, it is questionable whether gauge fields or fermion fields, both unobservable, should be required to satisfy this axiom. But to make sure that observables constructed from these unobservable fields conform with it, one may retain it, although it may be one condition too many as it is possibly the case for QED (see Section 2.4).

However, it should not be mistaken for the CCR or CAR (Section 1.6). Note that (2.2.9) is an operator identity which does not say anything about the case when the supports of f and g are not spacelike separated. We know that for a single free scalar field, this commutator is the distribution

$$[\varphi(f), \varphi(g)]_- = \Delta_+(f,g) - \Delta_+(g,f) \qquad \text{(free field case).} \tag{2.2.10}$$

It is interesting to see what happens if one assumes that the commutator of a generic scalar field yields a c-number, a case investigated by Greenberg [**Gre61**]: one can show that

- locality is implied by Poincaré invariance of the commutator, in turn a consequence of Poincaré covariance of the field (easy exercise);
- the field can be decomposed into a positive and a negative energy piece.

A field with this property has therefore been named *generalised free field* (see [**Stro93**] for a concise treatment).

2.2.2. Asymptotic fields. The Wightman axioms do not include the *condition of asymptotic completeness.* This essentially means that the field algebra $\mathcal{A}(\mathbb{M})$ contains elements which approach free fields in the limits $t \to \pm\infty$ and that the states these *asymptotic fields* create when applied to the vacuum fill up a dense subspace in the Hilbert space. Ruelle proved in [**Rue62**] that the above axioms imply the existence of asymptotic states if the theory has a *mass gap* and Buchholz succeeded in proving the massless case [**Bu75, Bu77**]. But *the existence of asymptotic states and fields does not imply asymptotic completeness* which is often written as

$$\mathfrak{H}_{in} = \mathfrak{H} = \mathfrak{H}_{out}, \tag{2.2.11}$$

where $\mathfrak{H}_{in}$ and $\mathfrak{H}_{out}$ are the Hilbert spaces of the incoming and outgoing particles. But *if* asymptotic completeness is given, the existence of a unitary S-matrix is guaranteed. We have already mentioned in Section 1.4 that the existence of the S-matrix has been proven for the superrenormalisable class $P(\varphi)_2$.

Nevertheless, as there is no compelling evidence for asymptotic completeness, Streater and Wightman decided to withdraw this condition from their list of axioms (Axiom IV in [**StreatWi00**], p.102).

2.2.3. Wightman distributions. The axioms translate directly to a package of properties of the vacuum expectation values. Let us now for simplicity confine ourselves to a single scalar field. It is not difficult to prove that the Wightman distributions defined by

$$W_n(f_1, ..., f_n) := \langle \Psi_0 | \varphi(f_1)...\varphi(f_n)\Psi_0 \rangle \tag{2.2.12}$$

have the following properties:

W1: Poincaré invariance. $W_n(f_1, ..., f_n) = W_n(\{a, \Lambda\}f_1, ..., \{a, \Lambda\}f_n)$ for all Poincaré transformations $(a, \Lambda) \in \mathscr{P}_+^\uparrow$. This is a simple consequence of the field's Poincaré covariance.

W2: Spectral condition. W_n vanishes if one test function's Fourier transform has its support outside the forward light cone, that is, if there is a j such that $\widetilde{f}_j(p) = 0$ for all $p \in \overline{V}_+$, then

$$\widetilde{W}_n(\widetilde{f}_1, ..., \widetilde{f}_n) = W_n(f_1, ..., f_n) = 0 \tag{2.2.13}$$

In this case one says that $\widetilde{W}_n$ and W_n have support inside the forward light cone $(\overline{V}_+)^n$. This property is a consequence of the spectral condition imposed by Axiom I.

But notice what it entails. While this is all very well for free fields, it raises serious doubts in a general QFT. If we just take the renormalised propagator of a scalar field in momentum space,

$$\widetilde{G}_r(p) = \frac{i}{p^2 - m_r^2 - \Sigma_r(p) + i0^+} = \lim_{\epsilon \downarrow 0} \frac{i}{p^2 - m_r^2 - \Sigma_r(p) + i\epsilon} \tag{2.2.14}$$

with physical mass $m_r > 0$ and self-energy $\Sigma_r(p)$, we have to ask ourselves whether this thing can actually do us a favour and vanish for spacelike momenta. In the case of a free field, this is well-understood as the integration over the zeroth component picks up the on-shell particles. However, let us assume that this mechanism also works for the distribution in (2.2.14).

But what about photons? This would mean that spacelike, ie t-channel photons effectively do not contribute to the two-point function (cf. Section 6.6). We already see here, the Wightman framework does not accommodate the Maxwell field in its edifice as straightforwardly and clearly as one might wish for!

W3: Hermiticity. $W_n(f_1, ..., f_n) = W_n(f_n^*, ..., f_1^*)^*$. This follows from $\varphi(f)^\dagger = \varphi(f^*)$.

W4: Causality/Locality. If f_j and f_{j+1} have mutually spacelike separated support, then

$$W_n(f_1, ..., f_j, f_{j+1}, ..., f_n) = W_n(f_1, ..., f_{j+1}, f_j, ..., f_n). \tag{2.2.15}$$

W5: Positivity. $\Psi_f = f_0\Psi_0 + \sum_{n\geq 1} \varphi(f_{n,1})...\varphi(f_{n,n})\Psi_0$ is the form of a general state in $\mathfrak{D}_0$. The property

$$\sum_{n\geq 0} \sum_{j+k=n} W_n(f_{j,j}^*, ..., f_{j,1}^*, f_{k,1}, ..., f_{k,k}) \geq 0 \tag{2.2.16}$$

is a consequence of the requirement $\langle \Psi_f | \Psi_f \rangle = ||\Psi_f||^2 \geq 0$, where $W_0 = |f_0|^2 \geq 0$.

W6: Cluster decomposition. Let $a \in \mathbb{M}$ be spacelike. Then

$$\lim_{\lambda \to \infty} W_n((f_1, ..., f_j, \{\lambda a, 1\}f_{j+1}, ..., \{\lambda a, 1\}f_n) = W_j(f_1, ..., f_j)W_{n-j}(f_{j+1}, ..., f_n). \tag{2.2.17}$$

Distributions with these features are called *Wightman distributions* because a given Wightman field satisfying the above axioms gives rise to such distributions.

What if one is handed a set of such distributions without any further information, is there a field theory with such distributions?

2.2.4. Reconstructing a quantum field theory. The reconstruction theorem asserts just that, up to unitary equivalence: every given set of Wightman distributions is associated to an existent Wightman field theory. A tentative first version of this result was published in [**Wi56**] when, as we have already pointed out, the axioms had not yet been formulated as such. It became a proper theorem once the axioms had been formulated.

Before we start let us have a look at an important result known as Schwartz's nuclear theorem [**StreatWi00**]. It states that for every multilinear tempered distribution $W : \mathscr{S}(\mathbb{M})^l \to \mathbb{C}$ there exists a tempered distribution S on $\mathscr{S}(\mathbb{M}^l)$ such that

$$W(f_1, ..., f_l) = S(f_1 \otimes ... \otimes f_l). \tag{2.2.18}$$

We remind the reader that the l-fold tensor product of functions $f_1, ..., f_l \in \mathscr{S}(\mathbb{M})$ is given by the function

$$(f_1 \otimes ... \otimes f_l)(x_1, ..., x_l) := f_1(x_1)...f_l(x_l), \tag{2.2.19}$$

which is an element in $\mathscr{S}(\mathbb{M}^l)$. This means practically that one can formally write

$$W(f_1, ..., f_n) = \int d^d x_1 \; ... \int d^d x_n \; \mathcal{W}(x_1, ..., x_n) f_1(x_1)...f_n(x_n) \tag{2.2.20}$$

with d being the dimension of spacetime $\mathbb{M}$. We will identify both distributions and simply write $W(f_1, ..., f_l) = W(f_1 \otimes ... \otimes f_l)$ as is customary. Here is the theorem.

Theorem 2.2 (Reconstruction theorem). *Let $\{W_n\}$ be a family of tempered distributions adhering to the above list of properties W1-W6. Then there is a scalar field theory fulfilling the Wightman axioms 0 to IV. Any other theory is unitarily equivalent.*

Proof. See Appendix Section B.2 □

The proof is constructive. We sketch it briefly. The underlying concept is the so-called *Borchers algebra* [**Bor62**]. It is given by the vector space

$$\mathscr{B} = \bigoplus_{n \geq 0} \mathscr{S}(\mathbb{M}^n) \tag{2.2.21}$$

of terminating sequences $(f_0, f_1, ..., f_n, 0, 0, ...)$ with $f_j \in \mathscr{S}(\mathbb{M}^j)$ and $\mathscr{S}(\mathbb{M}^0) := \mathbb{C}$. To make this space into an algebra, one defines the product by

$$f \times h := (f_0 h_0, f_0 \otimes h_1 + f_1 \otimes h_0, ..., (f \times g)_n, ...) \tag{2.2.22}$$

in which $f_0 \otimes h_0 = f_0 h_0$ is just the product in $\mathbb{C}$ and

$$(f \times h)_n := \sum_{j+k=n} [f_j \otimes g_k] \tag{2.2.23}$$

is the n-th component of the product. The Wightman distributions W_n are now represented by a so-called *Wightman functional* W on $\mathscr{B}$, given by

$$W(f \times h) := \sum_{n \geq 0} \sum_{j+k=n} W_n(f_j^* \otimes h_k), \tag{2.2.24}$$

where $W_0(f_0^* \otimes h_0) = f_0^* h_0$. Using this functional, one defines an inner product on $\mathscr{B}$ by setting $\langle f, h \rangle := W(f \times h)$ and makes it into a Hilbert space by the standard procedures of completion with respect to Cauchy sequences and nullifying zero norm states. A quantum field is then defined by the assignment of a Schwartz function $h \in \mathscr{S}(\mathbb{M})$ to the operator $\varphi(h)$ on the Hilbert space $\mathscr{B}$, given through the multiplication of a vector $g = (g_0, g_1, g_2, ...) \in \mathscr{B}$ by the vector $(0, h, 0, ...) \in \mathscr{B}$, ie the operator $\varphi(h)$ is the assigment

$$g \mapsto \varphi(h)g := (0, h, 0, ...) \times g = (0, h \otimes g_0, ..., h \otimes g_{n-1}, ...), \tag{2.2.25}$$

ie $(\varphi(h)g)_n = h \otimes g_{n-1}$ is the n-th component. The reader is referred to Appendix Section B.2 for a complete account of the proof.

2.3. Proof of Haag's theorem

As we shall see in this section, Haag's theorem relies strongly on the analyticity properties of the Wightman distributions (2.2.20). Although their Schwartz kernels $\mathcal{W}_n(x_1, ..., x_n)$ are generalised functions, the idea is to see them as the boundary values of meromorphic and hence locally holomorphic functions in the sense of distribution theory. In fact, this is the assertion of a theorem: any distribution in $\mathscr{S}(\mathbb{M}^n)$ is the boundary value of a meromorphic function (see [**Scho08**], Section 8.5). One speaks of these meromorphic functions as analytic continuations of the corresponding distributions.

Before we present Haag's theorem, we have to first go through a small battery of results used in its proof: the edge-of-the-wedge theorem, the Reeh-Schlieder theorem, a corollary of it and the theorem of Jost and Schroer.

2.3.1. Analytic continuation of tempered distributions. A pedagogical example of the analytic continuation of a tempered distribution is the meromorphic function

$$F(x + iy) = \frac{1}{x + iy} = \frac{x}{x^2 + y^2} - i\pi \frac{y}{\pi(x^2 + y^2)}. \tag{2.3.1}$$

It is perfectly holomorphic off the origin and gives rise to the tempered distribution

$$F(f) = \lim_{y \to 0} \int_{-\infty}^{+\infty} dx \; F(x + iy) f(x) = \mathcal{P} \int_{-\infty}^{+\infty} \frac{dx}{x} f(x) - i\pi f(0), \tag{2.3.2}$$

where $\mathcal{P} \int$ is the Cauchy principle value integral.

In the case of Wightman distributions, one can specify distinctly what values the imaginary parts are permitted to take. We roughly follow [**Stro13**]. If we consider the two-point Wightman function

$$\mathcal{W}(x - iy) = \int \frac{d^4 q}{(2\pi)^4} \; e^{-iq\cdot(x-iy)} \; \widetilde{W}(q), \tag{2.3.3}$$

we see that $q \cdot y > 0$ is required. Given the spectral property $\widetilde{W}(q) = 0$ if $q \notin \overline{V}_+$, then $y \in V_+$ suffices. This motivates the definition of the *forward tube*

$$\mathcal{T}_n := \mathbb{M}^n - i(V_+)^n. \tag{2.3.4}$$

For the two-point Wightman distribution we then pick $\eta \in V_+$ and get the identity

$$W(f, h) = \lim_{t \downarrow 0} \int d^4 x_1 \int d^4 x_2 \; f(x_1) \mathcal{W}(x_1 - x_2 - it\eta) h(x_2). \tag{2.3.5}$$

The *extended forward tube* is defined by $\mathcal{T}_n' := L_+(\mathbb{C})\mathcal{T}_n$, where $L_+(\mathbb{C})$ is the connected complex Lorentz group which is applied to each argument separately, ie

$$\mathcal{T}_n \ni z = (z_1, ..., z_n) \mapsto \Lambda z := (\Lambda z_1, ..., \Lambda z_n) \in \mathcal{T}_n'. \tag{2.3.6}$$

These transformations are defined by the two properties $\det(\Lambda) = 1$ and the invariance property

$$(\Lambda z) \cdot (\Lambda z) = z^2 = x^2 - y^2 + i \; 2x \cdot y, \tag{2.3.7}$$

where all multiplications are understood as Minkowski products and $x, y \in \mathbb{M}$. The kernel function can now be analytically continued to this extended tube by setting

$$\mathcal{W}(\Lambda z) := \mathcal{W}(z) \qquad \forall z \in \mathcal{T}_n. \tag{2.3.8}$$

Note that this is an unambiguous definition because different preimages are Lorentz-equivalent: let $u = \Lambda w = \Lambda' v$. Then

$$\mathcal{W}(u) = \mathcal{W}(w) = \mathcal{W}(\Lambda^{-1}\Lambda' v) = \mathcal{W}(v) \tag{2.3.9}$$

because $\Lambda^{-1}\Lambda' \in L_+(\mathbb{C})$. This is known as the *theorem of Bargmann, Hall and Wightman* (see [**Jo65**], Chapter 4). Note that because there is a complex Lorentz transformation such that $\Lambda z = -z$, the Wightman distributions are now also analytically continued to the *backward tube* $\mathbb{M}^n + i(V_+)^n$.

Jost points. However, the forward tube does not contain any real points because the imaginary parts of the complex arguments lie in $(V_+)^n$, ie $z \in \mathcal{T}_n$ implies $\Im(z_j) \neq 0$ for all $j = 1, ..., n$. An important property of the extended tube $\mathcal{T}_n'$ is that in contrast to $\mathcal{T}_n$, it *does* contain real points, the so-called *Jost points*, for which $\Im(z) = \Im((z_1, ..., z_n)) = 0$. Let us grant them a definition.

DEFINITION 2.3 (Jost points). *A point $z = (z_1, ..., z_n) \in \mathcal{T}_n' \subset \mathbb{C}^{4n}$ in the extended forward tube with vanishing imaginary part, that is, $\Im(z) = 0$, is called* Jost point. *By*

$$\mathrm{co}(z) = \left\{ \textstyle\sum_{j=1}^{n} \alpha_j z_j \;\middle|\; \textstyle\sum_{j=1}^{n} \alpha_j = 1, \forall j : \alpha_j \geq 0 \right\} \subset \mathbb{C}^4 \tag{2.3.10}$$

we denote the convex hull of a point $z \in \mathbb{C}^{4n}$.

Note that the convex hull of a point $z \in \mathbb{C}^{4n}$ lies in $\mathbb{C}^4$ for any $n \geq 1$. The following theorem due to Jost characterises Jost points and their convex hull.

THEOREM 2.4 (Jost). *Let $z = (z_1, ..., z_n) \in \mathcal{T}_n'$ be a Jost point. Then the convex hull of this point,* $\mathrm{co}(z) \subset \mathbb{C}^4$*, is spacelike, ie consists of spacelike points only:*

$$w \in \mathrm{co}(z) \quad \Rightarrow \quad w^2 < 0. \tag{2.3.11}$$

Conversely, a set of vectors $z_1, ..., z_n \in \mathbb{M}$ with spacelike convex hull comprise a Jost point in the extended forward tube $\mathcal{T}_n'$.

PROOF. We only prove the case $n = 1$ (for general $n \in \mathbb{N}$ see [**StreatWi00**], pp.70,71). If $z' \in \mathbb{M}$ is spacelike, then we can in a first step Lorentz-transform it to $z' = (z_0', z_1', 0, 0)$ with $z_1' > |z_0'|$. Next, we apply the complex Lorentz transformation

$$\begin{pmatrix} z_0 \\ z_1 \end{pmatrix} = \underbrace{\begin{pmatrix} \cos\alpha & i\sin\alpha \\ i\sin\alpha & \cos\alpha \end{pmatrix}}_{\Lambda \in L_+(\mathbb{C})} \begin{pmatrix} z_0' \\ z_1' \end{pmatrix} \tag{2.3.12}$$

for $\sin\alpha > 0$. Then $\Im(z) \in V_+$, ie $z \in \mathcal{T}_1$, to be verified by the reader. This means in particular $z' = \Lambda^{-1} z \in \mathcal{T}_1'$. Hence any spacelike vector in $\mathbb{M}$ is a Jost point in the extended tube $\mathcal{T}_1'$, ie the set of Jost points is not empty and contains all spacelike points for $n = 1$. Now let $z' \in \mathcal{T}_1'$ be a Jost point. Then there must be a complex Lorentz transformation $\Lambda \in L_+(\mathbb{C})$ and a vector $z \in \mathcal{T}_1$ such that $z' = \Lambda z$ (otherwise z' would not be in $\mathcal{T}_1'$). We write $z = x + iy$ and compute

$$(z')^2 = (\Lambda z) \cdot (\Lambda z) = z^2 = x^2 - y^2 + i2x \cdot y. \tag{2.3.13}$$

Because $\Im(z') = 0$ by assumption, we have $x \cdot y = 0$ which implies $x^2 < 0$ because $y \in V_+$. Thus, $(z')^2 = z^2 = x^2 - y^2 < 0$ and therefore all Jost points are spacelike. In the case $n > 1$, only spacelike vectors whose convex hull is spacelike form a Jost point. □

This has a very useful consequence for the two-point function because it depends only on one spacetime variable $\xi \in \mathbb{M}$: if one knows its values on the spacelike double cone, ie the set $\{\xi \in \mathbb{M} : \xi^2 < 0\}$ which, as Jost's theorem informs us, habours all Jost points, then there is an open subset $\mathcal{O} \subset (\mathbb{M} + iV_+) \cup (\mathbb{M} - iV_+)$ in the forward tube and the backward tube, to which it can be analytically continued by applying the complex Lorentz group $L_+(\mathbb{C})$. The so-called 'edge-of-the-wedge' theorem, to be presented next, then says that the two-point function is uniquely characterised there.

For the higher-point functions, one needs more subtle arguments to show that the Jost points comprise a subset large enough to uniquely characterise them (see [**StreatWi00**], pp.70,71).

2.3.2. Edge of the wedge. The next result, known as 'edge-of-the-wedge' theorem, is crucial for Haag's theorem because it guarantees in particular that the two-point function is sufficiently characterised on the spacelike points of $\mathbb{M}$, ie the Jost points of $\mathcal{T}_1' \subset \mathbb{C}^4$.

THEOREM 2.5 (Edge of the wedge). *Let $\mathcal{O} \subset \mathbb{C}^{4n}$ be an open subset which contains a real open subset $E \subset \mathcal{O}$. Suppose $F_\pm$ is holomorphic in $D_\pm := [\mathbb{M}^n \pm i(V_+)^n] \cap \mathcal{O}$ and for any $y \in V_+$ one finds*

$$\lim_{t \downarrow 0} F_+(x + ity) = \lim_{t \downarrow 0} F_-(x - ity) \qquad \forall x \in E \tag{2.3.14}$$

in the sense of distributions. Then there is a function G holomorphic in an open complex neighbourhood N of E such that $G = F_\pm$ on $D_\pm$.

PROOF. See [**StreatWi00**], Theorem 2-16. □

In the case of the two-point function $\mathcal{W}(\xi)$ one starts with $E = \{\xi \in \mathbb{M} : \xi^2 < 0\}$, ie the set of all spacelike vectors. Any complex open neighbourhood $N \subset \mathbb{C}^4$ of E must contain both lightlike and timelike vectors in $\overline{V}_+$ and $-\overline{V}_+$. From there, it is clear that by following the Lorentz group's orbits, $\mathcal{W}(\xi)$ is given everywhere in $\overline{V}_+ \cup (-\overline{V}_+)$.

2.3.3. Local operator algebras. The following result, called Reeh-Schlieder theorem, is interesting from a physical point of view despite its mathematical fancyness. It decribes the remarkable fact that for any open $E \subset \mathbb{M}$, the local operator algebra

$$\mathcal{A}(E) = \langle\, \varphi(f), \varphi(f)^\dagger : f \in \mathscr{D}(E) \,\rangle_{\mathbb{C}} \tag{2.3.15}$$

is cyclic for the vacuum, ie $\mathcal{A}(E)\Psi_0 \subset \mathfrak{H}$ is dense, no matter how small $E \subset \mathbb{M}$ is!

Let us say we choose the open compact subset $E \subset \mathbb{M}$ to be very tiny. The theorem tells us that every state can be generated from the corresponding local field operators, ie the operators $\varphi(f)$ smeared with $f \in \mathscr{D}(E)$. This means physically that if we know what happens in the tiny region E of spacetime $\mathbb{M}$, we know what may happen anywhere.

In high energy physics, this makes sense: to find out what happens in a tiny subset E, we need to use a gargantuan amount of energy. Then, as we know from experiments, more particles (or particle species for more field types) will show up. This suffices to be informed about what might occur anywhere in the universe under the same circumstances. Therefore, this result is not unphysical.

THEOREM 2.6 (Reeh-Schlieder). *Let $E \subset \mathbb{M}$ be open. If Ψ_0 is cyclic for $\mathcal{A}(\mathbb{M})$, then it is also cyclic for $\mathcal{A}(E)$. This means in particular that vectors of the form*

$$\Psi_f = f_0\Psi_0 + \sum_{j=1}^{n} \varphi(f_{j,1})...\varphi(f_{j,j})\Psi_0 \qquad \textit{('localised states')} \tag{2.3.16}$$

with $f_0 \in \mathbb{C}$ and $f_{j,1}, ..., f_{j,j} \in \mathscr{D}(E)$ for all $j = 1, ..., n$ are dense in $\mathfrak{H}$.

PROOF. We follow [**StreatWi00**]. Let $\Psi \in \mathfrak{H}$ be orthogonal to all vectors in $\mathcal{A}(E)\Psi_0$, then

$$F(f_1, ..., f_n) = \langle \Psi | \varphi(f_1)...\varphi(f_n)\Psi_0 \rangle \tag{2.3.17}$$

is a translation-invariant distribution vanishing for all $f_1, ..., f_n \in \mathscr{D}(E)$. By the nuclear theorem, we may write

$$F(f_1, ..., f_n) = F(f_1 \otimes ... \otimes f_n). \tag{2.3.18}$$

Then, for any $\eta \in (V_+)^{n-1}$ there exists a meromorphic function $\mathcal{F}_n$ such that

$$F(f_1 \otimes ... \otimes f_n) = \lim_{t \downarrow 0} \int d^d x_1 \, ... \int d^d x_n \; \mathcal{F}_n(\xi - it\eta) f_1(x_1)...f_n(x_n), \tag{2.3.19}$$

where $\xi = (\xi_1, ..., \xi_{n-1})$ are the variables $\xi_j := x_j - x_{j+1}$, $j = 1, ..., n-1$. The limit exists by virtue of the spectral property (Axiom I). By assumption we know, however, that (2.3.19) vanishes for $\xi \in \mathbb{M}^{n-1}$ with the property $(x_1, ..., x_n) \in E^n$, ie

$$\lim_{t \downarrow 0} \mathcal{F}_n(\xi - it\eta) = 0 \tag{2.3.20}$$

in the sense of distribution theory for all $f_j \in \mathscr{D}(E)$. By Theorem 2.5 (edge-of-the-wedge) we conclude that F vanishes on all elements $f_1 \otimes ... \otimes f_n \in \mathscr{S}(\mathbb{M})^{\otimes n}$: the argument in that theorem can be iterated to reach any point in $\mathbb{M}$. But this means $\langle \Psi | \mathcal{A}(\mathbb{M}) \Psi_0 \rangle = 0$ and thus $\Psi = 0$, on account of $\mathfrak{D}_0 = \mathcal{A}(\mathbb{M})\Psi_0$ being dense, guaranteed by Axiom II. This entails that $\mathcal{A}(E)\Psi_0$ is dense in $\mathfrak{H}$ because we have found that a vector $\Psi \in \mathfrak{H}$ orthogonal to this set must vanish. $\square$

This theorem has an important consequence. Let $E' := \{x \in \mathbb{M} | (x-y)^2 < 0 \; \forall y \in E\}$ be the spacelike complement of E. Then there exists no annihilating operator in $\mathcal{A}(E')$.

Corollary 2.7. *Let $E \subset \mathbb{M}$ be open and $T \in \mathcal{A}(E')$ such that $T\Psi_0 = 0$. Then $T = 0$ weakly, ie there exists no annihilator in $\mathcal{A}(E')$.*

Proof. Pick any $\Psi \in \mathcal{A}(E)\Psi_0$ and let $P \in \mathcal{A}(E)$ be such that $\Psi = P\Psi_0$. For any $\Phi \in \mathfrak{D}$ in the domain of the field operators we consider

$$\langle \Psi | T^\dagger \Phi \rangle = \langle T\Psi | \Phi \rangle = \langle TP\Psi_0 | \Phi \rangle = \langle PT\Psi_0 | \Phi \rangle = 0 \tag{2.3.21}$$

since, by virtue of the field's locality, one has $[P, T] = 0$ due to $P \in \mathcal{A}(E)$ and $T \in \mathcal{A}(E')$. Because $\mathcal{A}(E)\Psi_0$ is dense, $T^\dagger \Phi = 0$ follows for any $\Phi \in \mathfrak{D}$. Let now $\Psi \in \mathfrak{D}$, then

$$\langle T\Psi | \Phi \rangle = \langle \Psi | T^\dagger \Phi \rangle = 0 \qquad \forall \Phi \in \mathfrak{D} \tag{2.3.22}$$

entails $T\Psi = 0$ because $\mathfrak{D}$ is dense. As this holds for all $\Psi \in \mathfrak{D}$, the assertion follows because T vanishes weakly, ie $\langle T\Psi | \Phi \rangle = 0$ is true for all $\Psi, \Phi \in \mathfrak{D}$. $\square$

2.3.4. Haag's theorem II. The final piece we need for the proof of Haag's theorem is the so-called *Jost-Schroer theorem* which says that a theory whose two-point function coincides with that of a free field of mass $m > 0$ is itself such a theory.

Theorem 2.8 (Jost-Schroer Theorem). *Let φ be a scalar field whose two-point distribution coincides with that of a free field of mass $m > 0$, ie*

$$\langle \Psi_0 | \varphi(f) \varphi(h) \Psi_0 \rangle = \int \frac{d^4 p}{(2\pi)^4} \; \widetilde{f}^*(p) \widetilde{\Delta}_+(p; m^2) \widetilde{h}(p), \tag{2.3.23}$$

where $\widetilde{\Delta}_+(p; m^2) = 2\pi\theta(p_0)\delta(p^2 - m^2)$. Then φ is itself a free field of mass $m > 0$.

Proof. If we define a free scalar field through the condition $\varphi([\square + m^2]f) = 0$ for all Schwartz functions $f \in \mathscr{S}(\mathbb{M})$, then the proof is short and simple. First, note that (2.3.23) implies

$$\langle \Psi_0 | \varphi([\square + m^2]f) \varphi([\square + m^2]h) \Psi_0 \rangle = 0 \tag{2.3.24}$$

and thus $||\varphi([\square + m^2]f)\Psi_0|| = 0$, that is, $\varphi([\square + m^2]f)\Psi_0 = 0$ for any Schwartz function f. Then, by locality of the field (Axiom IV) and Corollary 2.7, $\varphi([\square + m^2]f) = 0$ and thus φ is a free field in this sense. The full proof shows that this property entails that all Wightman distributions coincide with those of a free field, see Appendix Section B.3. $\square$

Since, as we have already mentioned in Subsection 1.2.4, Pohlmeyer has proved this assertion for the massless case, the Jost-Schroer theorem holds for fields with mass $m \geq 0$ [**Po69**]. Now, finally, we shall see that the proof of Haag's theorem is implicitly contained in the package of the statements above together with the equality of the two-point functions (1.2.16) at equal times.

THEOREM 2.9 (Haag's Theorem). *Let φ and φ_0 be two Hermitian scalar fields of mass $m \geq 0$ in the sense of the Wightman framework. Suppose the sharp-time limits $\varphi(t, f)$ and $\varphi_0(t, f)$ exist and that at time $t = 0$ these two sharp-time fields form an irreducible set in their respective Hilbert spaces $\mathfrak{H}$ and $\mathfrak{H}_0$. Furthermore let there be an isomorphism $V : \mathfrak{H}_0 \to \mathfrak{H}$ such that at time $t = 0$*

$$\varphi(0, f) = V\varphi_0(0, f)V^{-1}. \tag{2.3.25}$$

Then φ is also a free field of mass $m \geq 0$.

PROOF. We start by showing that $U(\mathbf{a}, R) = VU_0(\mathbf{a}, R)V^{-1}$ for the two representations of the Euclidean subgroups in $\mathfrak{H}$ and $\mathfrak{H}_0$, respectively. We use covariance with respect to the Euclidean subgroup (Axiom III) and unitary equivalence (2.3.25):

$$\begin{aligned}
&\varphi(0, f)U(\mathbf{a}, R)^\dagger VU_0(\mathbf{a}, R)V^{-1} = U(\mathbf{a}, R)^\dagger\varphi(0, \{\mathbf{a}, R\}f)VU_0(\mathbf{a}, R)V^{-1}\\
&= U(\mathbf{a}, R)^\dagger V\varphi_0(0, \{\mathbf{a}, R\}f)U_0(\mathbf{a}, R)V^{-1} = U(\mathbf{a}, R)^\dagger VU_0(\mathbf{a}, R)\varphi_0(0, f)V^{-1}\\
&= U(\mathbf{a}, R)^\dagger VU_0(\mathbf{a}, R)V^{-1}\varphi(0, f).
\end{aligned} \tag{2.3.26}$$

On account of the irreducibility of φ's field algebra, warranted by Axiom II, we have

$$U(\mathbf{a}, R)^\dagger VU_0(\mathbf{a}, R)V^{-1} = c(\mathbf{a}, R)1 \tag{2.3.27}$$

for some $c(\mathbf{a}, R) \in \mathbb{C}$ which must be constant and, in fact, equal to 1 due to the group property and unitarity of the representation (Axiom O). Thus, $U(\mathbf{a}, R)V = VU_0(\mathbf{a}, R)$. Let Ω and Ω_0 denote the two vacua. Then

$$U(\mathbf{a}, R)V\Omega_0 = VU_0(\mathbf{a}, R)\Omega_0 = V\Omega_0, \tag{2.3.28}$$

which means $V\Omega_0 = a\Omega$. By unitarity of V we have $|a| = 1$. We are allowed to set $a = 1$ (or absorb a into the definition of V). The consequence of this is

$$\langle\Omega_0|\varphi_0(0, f)\varphi_0(0, h)\Omega_0\rangle = \langle\Omega_0|V^{-1}\varphi(0, f)VV^{-1}\varphi(0, h)V\Omega_0\rangle = \langle\Omega|\varphi(0, f)\varphi(0, h)\Omega\rangle \tag{2.3.29}$$

and hence $\langle\Omega_0|\varphi_0(0, \mathbf{x})\varphi_0(0, \mathbf{y})\Omega_0\rangle = \langle\Omega|\varphi(0, \mathbf{x})\varphi(0, \mathbf{y})\Omega\rangle$ in the sense of distributions. The only singular point is where $\mathbf{x} = \mathbf{y}$, as we know because both expressions are the well-known two-point function of a free scalar field of mass $m \geq 0$ at equal times, ie

$$\Delta_+(0, \mathbf{x} - \mathbf{y}; m^2) = \langle\Omega_0|\varphi_0(0, \mathbf{x})\varphi_0(0, \mathbf{y})\Omega_0\rangle = \langle\Omega|\varphi(0, \mathbf{x})\varphi(0, \mathbf{y})\Omega\rangle. \tag{2.3.30}$$

For any spacelike point $x = (t, \xi)$, one can find a Lorentz transformation $\Lambda \in \mathscr{L}_+^\uparrow$ that takes it into the zero-time slice: $\Lambda x = x' = (0, \xi')$. Therefore, we have by Lorentz invariance (implied by Axiom III)

$$\Delta_+(x - y; m^2) = \langle\Omega_0|\varphi_0(x)\varphi_0(y)\Omega_0\rangle = \langle\Omega|\varphi(x)\varphi(y)\Omega\rangle. \tag{2.3.31}$$

for $(x - y)^2 < 0$. By the edge-of-the-wedge Theorem 2.5, this means the two-point Wightman distribution of the field φ agrees with that of the free field. Thus, on account of the Jost-Schroer Theorem 2.8 in case $m > 0$ and Pohlmeyer's version for $m = 0$, φ is a free field of mass $m \geq 0$. $\square$

2.3.5. Summary: provisos of the proof. As the exposition shows, the proof of Haag's theorem makes use of all Wightman axioms and makes two additional strong assumptions:

(1) the existence of sharp-time fields $\varphi_0(t,f), \varphi(t,f)$ forming an irreducible operator algebra at $t=0$ and
(2) unitary equivalence between both fields.

Here is a summary of the salient stages of the proof, the provisos they rely on and their consequences.

- Poincaré-invariance of both vacua (Axiom O), Poincaré covariance of both fields (Axiom III) and their unitary equivalence (2.3.25) jointly entail that the two-point functions of both theories agree for spacelike separated points,

$$\langle\Omega_0|\varphi_0(x)\varphi_0(y)\Omega_0\rangle = \langle\Omega|\varphi(x)\varphi(y)\Omega\rangle \qquad \text{if} \quad (x-y)\in E, \tag{2.3.32}$$

  where $E=\{\xi\in\mathbb{M}:\xi^2<0\}$ is the spacelike double cone.
- Jost's theorem (Theorem 2.4) tells us that E is comprised of Jost points which are the real points of the extended forward tube $\mathcal{T}_1'$. This set is in turn the image of the complex Lorentz group when applied to the forward tube $\mathcal{T}_1=\mathbb{M}-iV_+$. This means that starting from the spacelike double cone E lying in $\mathcal{T}_1'$, the two-point functions can be analytically continued into the forward tube $\mathcal{T}_1$ where they constitute an open set $\mathcal{O}\subset\mathcal{T}_1\subset\mathbb{C}^4$. Axioms O-III are needed since these objects must be Poincaré-invariant distributions that satisfy the spectral property (guaranteed by Axiom I). The edge-of-the-wedge theorem makes sure that the analytic continuations of both two-point functions really agree, ie

$$\langle\Omega_0|\varphi_0(x)\varphi_0(y)\Omega_0\rangle = \langle\Omega|\varphi(x)\varphi(y)\Omega\rangle \qquad \text{for all} \quad (x-y)\in\mathbb{M}. \tag{2.3.33}$$

- Given this result, the Jost-Schroer theorem can now be put in place. It relies on the Reeh-Schlieder Theorem 2.6 and its Corollary 2.7. For these theorems to be applicable, the spectral property of the Wightman distributions (Axioms O-III) and locality (Axiom IV) must be fulfilled.

So we see that *all Wightman axioms must be fulfilled* for the proof of Haag's theorem. No use is made of the conjugate momentum field $\pi(f)$. Many proofs including the one in [**StreatWi00**] use this field in addition and impose the intertwining relation

$$\pi(0,f)=V\pi_0(0,f)V^{-1}. \tag{2.3.34}$$

However, it is not employed in any of the proof's steps except implicitly when irreducibility is used to obtain the relation between the two Poincaré representations in (2.3.27).

The fact is, the free field φ_0 itself is a fully fledged irreducible Wightman field without its conjugate field: it generates a dense subspace in $\mathfrak{H}_0$ all by itself even at one fixed time $t=0$! Irreduciblity of $\varphi_0(0,f)$ then follows from cyclicity, as already mentioned (see Axiom II). So the difference between our proof, which agrees with Roman's version in Section 8.4 of his textbook [**Ro69**], and Wightman's lies in the assumption of what constitutes an irreducible field algebra.

Because the Wightman axioms do not include the CCR and the vacuum is cyclic for the free field which then, as a consequence, is irreducible, the conjugate momentum field need not join the game. It only brings in an additional strong assumption about the existence of the time derivative of the field φ.

2.4. Haag's theorem for fermion and gauge fields

At first glance, Haag's theorem in the above form holds only for scalar fields. It can, with some notational inconvenience, also be formulated for a collection $\{\varphi_j\}$ of scalar fields. But for fermion and gauge fields whose Poincaré covariance is nontrivial, things seem to look slightly different. Especially the Jost-Schroer theorem does not appear to carry over directly. We shall briefly see in this section that while all arguments used in the proof of Haag's theorem go also through for fermion fields, this turns out to be naive for gauge theories.

So assume that the Wightman framework with Axioms O - IV is true for a field with spin $s > 0$, then the equality of both theories' n-point functions on spacelike separated points is still given for $n \leq 4$. This is the assertion labelled 'generalised Haag's theorem' by Streater and Wightman in [**StreatWi00**].

We shall briefly go through its proof and show that because the Jost-Schroer theorem also holds for anticommuting fields, Haag's theorem affects fields of nontrivial spin ($s > 0$) as well. Note that it is irrelevant how many different collections of fields we put into both Hilbert spaces as long as they are connected via the same intertwiner V.

2.4.1. Dirac fields. Let $\psi_j(t, \mathbf{x})$ and $\psi_j^0(t, \mathbf{x})$ be two Dirac fields in Wightman's sense, the latter field being free. We denote their respective vacua as Ω and Ω_0. We start with

$$\langle \Omega | \psi_j(0, \mathbf{x}) \overline{\psi}_l(0, \mathbf{y}) \Omega \rangle = \langle \Omega_0 | \psi_j^0(0, \mathbf{x}) \overline{\psi}_l^0(0, \mathbf{y}) \Omega_0 \rangle, \tag{2.4.1}$$

in which $\overline{\psi} = \psi^\dagger \gamma^0$. If we perform a Lorentz transformation Λ such that $\Lambda(0, \mathbf{x}) = x$, $\Lambda(0, \mathbf{y}) = y$ and $(x - y)^2 < 0$ on both sides, we get

$$\sum_{a,b} S_{ja}(\Lambda^{-1}) S_{bl}(\Lambda^{-1})^* \langle \Omega | \psi_a(x) \overline{\psi}_b(y) \Omega \rangle = \sum_{a,b} S^0_{ja}(\Lambda^{-1}) S^0_{bl}(\Lambda^{-1})^* \langle \Omega_0 | \psi_a^0(x) \overline{\psi}_b^0(y) \Omega_0 \rangle, \tag{2.4.2}$$

with the corresponding spinor representations of the Lorentz group. If we strip the lhs of (2.4.2) of its spinor representation matrices, we obtain

$$\langle \Omega | \psi_j(x) \overline{\psi}_l(y) \Omega \rangle = \sum_{a,b} \sum_{a',b'} S_{ja'}(\Lambda) S^0_{a'a}(\Lambda^{-1}) S_{b'l}(\Lambda)^* S^0_{bb'}(\Lambda^{-1})^* \langle \Omega_0 | \psi_a^0(x) \overline{\psi}_b^0(y) \Omega_0 \rangle. \tag{2.4.3}$$

For the relation between the vacua $V\Omega_0 = \Omega$, derived along the same lines as in (2.3.26) to obtain (2.3.27), one chooses the same spinor representation for both fields. This is sensible and not a loss of generality because these representations are unitarily equivalent. Then follows for spacelike $\xi = x - y$

$$\langle \Omega | \psi_j(x) \overline{\psi}_l(y) \Omega \rangle = \langle \Omega_0 | \psi_j^0(x) \overline{\psi}_l^0(y) \Omega_0 \rangle. \tag{2.4.4}$$

It is equally valid for the n-point functions if $n \leq 4$. Analytic continuation and edge-of-the-wedge theorem tell us that (2.4.4) holds everywhere in spacetime $\mathbb{M}$. This concludes the proof of Streater and Wightman's generalised Haag's theorem.

Because the rhs is the two-point function of the free propagator, one finds that for

$$j_a(x) := \sum_b (i\gamma^\mu \partial_\mu - m)_{ab} \psi_b(x) \tag{2.4.5}$$

one has

$$\langle \Omega | j_a(x) \overline{j}_a(y) \Omega \rangle = 0 \tag{2.4.6}$$

and thus $||j_a^\dagger(f)\Omega|| = 0$ for every spinor index a and Schwartz function f. Since Corollary 2.7 is also true for anticommuting fields ([**StreatWi00**], remark on p.139) it follows $j_a^\dagger(f) = 0$. Therefore, we see that the Jost-Schroer theorem and consequently Haag's theorem hold also true for fermion fields.

2.4.2. Axioms and Haag's theorem for gauge theories. In contrast to what the previous paragraph suggests, gauge fields are different and prove a recusant species. As every physics student learns in a basic QFT course, the quantisation of the photon field is not as straightforward as for fermions and scalar bosons. If we take the Wightman axioms as an alternative quantisation programme for the photon field $A_\mu(x)$ and hence its field strength tensor $F_{\mu\nu}(x)$, then, as Strocchi has found, we end up with a very bad form of triviality [**Stro67, Stro70**].

THEOREM 2.10 (Strocchi). *Let $A_\mu(x)$ be an operator-valued distribution adhering to Axioms 0,I and II. Assume furthermore that that the field is Poincaré covariant (Axiom III) according to*

$$U(a,\Lambda)A_\mu(x)U(a,\Lambda)^\dagger = \Lambda^\sigma{}_\mu A_\sigma(\Lambda x + a) \tag{2.4.7}$$

or assume that Axiom IV is satisfied, ie locality $[A_\mu(x), A_\mu(x)] = 0$ for $(x-y)^2 < 0$. If the operator-valued distribution

$$F_{\mu\nu}(x) := \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x) \tag{2.4.8}$$

satisfies the free Maxwell equation $\partial_\mu F^{\mu\nu} = 0$, then

$$\langle\Omega|F_{\mu\nu}(x)F_{\rho\sigma}(y)\Omega\rangle = 0 \qquad \forall x,y \in \mathbb{M}, \tag{2.4.9}$$

where Ω is the vacuum. If both axioms III and IV are fulfilled, that is, all Wightman axioms, one finds $F_{\mu\nu}(x) = 0$ which means $A_\mu(x) = \partial_\mu\varphi(x)$, ie the photon field is a gradient field.

PROOF. We shall only sketch the proof for the case of Poincaré covariance (Axiom III). For a thorough exposition, the reader is referred to [**Stro67**] and [**Stro70**].

The first thing to show is that $D_{\mu\nu}(x,y) := \langle\Omega|A_\mu(x)A_\nu(y)\Omega\rangle$ is of the form

$$D_{\mu\nu}(\xi) = g_{\mu\nu}D_1(\xi) + \partial_\mu\partial_\nu D_2(\xi), \tag{2.4.10}$$

where $\xi = x - y$ and $D_j(\xi) = D_j(\Lambda\xi)$ for $j = 1,2$. This is familiar to physicists in momentum space. Strocchi invokes a theorem by Araki and Hepp [**He63**] to prove it. Because the gauge field obeys $[\delta^\mu_\nu\Box - \partial^\mu\partial_\nu]A_\mu(x) = 0$, one has

$$[\delta^\mu_\nu\Box - \partial^\mu\partial_\nu]D_{\mu\rho}(\xi) = 0. \tag{2.4.11}$$

and hence $[g_{\mu\nu}\Box - \partial_\mu\partial_\nu]D_1(\xi) = 0$. A lemma asserting that a Lorentz-invariant function $F(x)$ with

$$[g_{\mu\nu}\Box + \alpha\partial_\mu\partial_\nu]F(x) = 0\ , \qquad \alpha \notin \{0,4\} \tag{2.4.12}$$

must be constant incurs $D_1(\xi) = \text{const.}$ from which (2.4.9) follows. If locality comes on top, $F_{\mu\nu}(x) = 0$ is proven from $||F_{\mu\nu}(f)\Omega|| = 0$ by invoking the Jost-Schroer theorem. □

This result tells us that the photon field cannot reasonably be quantised by postulating the Wigthman axioms in the form given in Section 2.2. However, the situation can be remedied by abandoning the Maxwell equation $\partial_\mu F^{\mu\nu} = 0$ in favour of

$$\partial_\mu F^{\mu\nu} = -\partial^\nu(\partial_\mu A^\mu) \tag{2.4.13}$$

with the extra condition for the rhs' annihalators that $(\partial_\mu A^\mu)^+\mathfrak{H}' = 0$ on a subspace $\mathfrak{H}' \subset \mathfrak{H}$, as explained in [**StroWi74**]) which we shall draw on in the following.

When part of a quantisation scheme, this is known as the *Gupta-Bleuler condition*, as introduced in [**Gu50**] and [**Bleu50**]. The problem is that the inner product needed to define the Wightman distributions of QED is no longer definite, even on $\mathfrak{H}'$. Let us denote this inner product by $(\cdot,\cdot)$. On $\mathfrak{H}'$, one finds at least that it is positive semidefinite, ie

$$(\Psi,\Psi) \geq 0 \qquad \forall\Psi \in \mathfrak{H}', \tag{2.4.14}$$

but $\Phi \in \mathfrak{H}'$ with $(\Phi,\Phi) = 0$ does not mean Φ vanishes, it may well be a nonvanishing zero-norm state ($\Phi \neq 0$). We denote the subspace of such vectors by $\mathfrak{H}''$. Then, as a consequence of the Gupta-Bleuler condition, Maxwell's equations

$$(\Psi, \partial_\mu F^{\mu\nu}\Phi) = 0 \qquad \forall\Psi,\Phi \in \mathfrak{H}' \tag{2.4.15}$$

are fulfilled on $\mathfrak{H}'$. As a last step one obtains a physical Hilbert space $\mathfrak{H}_{\text{phys}}$ by taking the quotient $\mathfrak{H}_{\text{phys}} := \mathfrak{H}'/\mathfrak{H}''$ through nullifying all zero-norm states. In fact, the original vector space of states $\mathfrak{H}$ can be recovered as a Hilbert space provided one condition is met: there exists a sequence of seminorms (ρ_n) such that ρ_n is a seminorm on $\mathscr{S}(\mathbb{M}^n)$ and the Wightman distributions satisfy

$$|W_{n+m}(f_n^* \otimes g_m)| \leq \rho_n(f_n)\rho_m(g_m) \tag{2.4.16}$$

for any $f_n \in \mathscr{S}(\mathbb{M}^n)$ and $g_m \in \mathscr{S}(\mathbb{M}^m)$. Then there exist a bounded Hermitian operator η on $\mathfrak{H}$, called *metric operator*, and a (genuine) scalar product $\langle \cdot, \cdot \rangle$ such that

$$(\Phi, \Psi) = \langle \Phi, \eta\Psi \rangle \tag{2.4.17}$$

for all $\Phi, \Psi \in \mathfrak{D}$, where $\mathfrak{D}$ is the dense subspace generated by the field algebra. In words, the indefinite inner product $(\cdot, \cdot)$ of the Gupta-Bleuler formalism can be embedded into a Hilbert space through the employment of a sesquilinear form known as metric operator. This is in fact part of a reconstruction theorem of some sort for QED discussed by Yngvason in [**Yng77**] to which we refer the interested reader for details.

Note that the Wightman distributions of the gauge field A_μ and its field strength tensor $F_{\mu\nu}$ are given by the vaccum expectation values with respect to the indefinite inner product. On account of these complications, the Wightman axioms must be altered for gauge theories. As we have already explained in Section 2.2, it is Axiom O which needs to be modified: the Hilbert space is replaced by an inner product space $\mathfrak{H}$ whose inner product $(\cdot, \cdot)$ is nondegenerate. Because one can construct a Hilbert space structure from it, one has a so-called *Krein space* [**Stro93**]. So here is Axiom O for gauge theories [**Stei00**]:

- **Axiom O' (Relativistic complex vector space with nondegenerate form)**. The states of the physical system are described by (unit rays of) vectors in a separable complex vector space $\mathfrak{H}$ equipped with a nondegenerate form and a strongly continuous unitary representation $(a, \Lambda) \mapsto U(a, \Lambda)$ of the connected Poincaré group $\mathscr{P}_+^\uparrow$. Moreover, there is a unique state $\Psi_0 \in \mathfrak{H}$, called the *vacuum*, which is invariant under this representation, ie

$$U(a, \Lambda)\Psi_0 = \Psi_0 \qquad \text{for all } (a, \Lambda) \in \mathscr{P}_+^\uparrow. \tag{2.4.18}$$

To guarantee that a physical Hilbert space can be constructed, Axiom II must be augmented by Yngvason's condition (2.4.16), as proposed by Strocchi [**Stro93**]. One may then be content with these modified axioms for photon fields and prove Haag's theorem invoking the same arguments as we have in the case of Dirac fields above.

However, the following deliberations show that this contention cannot be maintained. The presented results reveal that any nonfree Maxwell theory is fundamentally at loggerheads with Wigthman's framework, probably the reason why Strocchi believes QED and nonabelian gauge theories are not afflicted by Haag's theorem [**Stro13**].

Maxwell's equations. To depart from free Maxwell theory and implement interactions with matter, one needs to introduce *charged fields.* We shall see now that in the presence of such fields, the last step to construct $\mathfrak{H}_{\text{phys}}$ leads to a triviality result. Therefore, as yet, *the last word on the Wightman axioms for QED has not been spoken.*

Before we jump to a conclusion about QED, let us survey the pertinent results obtained by Ferrari, Picasso and Strocchi [**FePStro74**]. A curious aspect of their exposition is that they make no reference to the gauge field A_μ and the corresponding Lagrangian formalism. Instead, they only use the field strength tensor $F_{\mu\nu}$, the charge current j^ν and a charged field ϕ which we will introduce now.

First of all, $F_{\mu\nu}$ and j_μ are operator-valued distributions, both local and relativistic in the sense of tensor fields, ie transforming under $\mathscr{P}_+^\uparrow$ according to

$$U(a, \Lambda)j_\mu(x)U(a, \Lambda)^\dagger = \Lambda^\sigma{}_\mu j_\sigma(\Lambda x+a) \ , \quad U(a, \Lambda)F_{\mu\nu}(x)U(a, \Lambda)^\dagger = \Lambda^\sigma{}_\mu \Lambda^\rho{}_\nu F_{\sigma\rho}(\Lambda x+a). \tag{2.4.19}$$

The field strength is supposed to be an antisymmetric tensor: $F_{\mu\nu} = -F_{\nu\mu}$. The charge current j^μ gives rise to a charge through its zeroth component

$$Q_R := j^0(f_u \otimes f_R) = \int d^4x \; f_u(x^0) f_R(\mathbf{x}) j^0(x) \tag{2.4.20}$$

where $f_R \in \mathscr{D}(\mathbb{R}^3)$ is such that $f_R(\mathbf{x}) = 1$ inside a ball of radius R, while vanishing rapidly outside the ball and f_u has compact support in $(-u, u) \subset \mathbb{R}$ with $\int_{\mathbb{R}} dx^0 f_u(x) = 1$. A scalar field ϕ is called *local relative to* j^μ in case it satisfies

$$[\phi(x), j^\mu(y)] = 0 \tag{2.4.21}$$

for spacelike distance, ie $(x-y)^2 < 0$. A field ϕ of this type is said to have charge q if

$$\lim_{R\to\infty} [Q_R, \phi(f)] = -q\phi(f) \tag{2.4.22}$$

for any $f \in \mathscr{D}(\mathbb{M})$. This introduces a global gauge transformation ('gauge transformation of the first kind'). Then, one has

THEOREM 2.11 (No Maxwell equations). *Let $j_\mu = \partial^\alpha F_{\alpha\mu}$ and ϕ be local relative to j_μ. Then*

$$\lim_{R\to\infty} [Q_R, \phi(f)] = 0, \tag{2.4.23}$$

ie the Maxwell equations cannot hold if the charge current is supposed to generate a nontrivial gauge transformation of the first kind (2.4.22).

PROOF. Sketch, for details see [**FePStro74**]. The main argument is that

$$[Q_R, \phi(f)] = [j^0(f_u \otimes f_R), \phi(f)] = -[F^{\mu 0}(\partial_\mu(f_u \otimes f_R)), \phi(f)] = -[F^{j0}(f_u \otimes \partial_j f_R), \phi(f)] \tag{2.4.24}$$

vanishes in the limit because f_R is constant in the region where the commutator receives contributions from the charge current integral (2.4.20), ie $\partial_j f_R = 0$ inside the ball. □

To cure this pathology, we abandon the strong form of Maxwell's equations in Theorem 2.11 and replace it by their weaker Gupta-Bleuler form

$$(\Psi, [\partial_\mu F^{\mu\nu}(f) - j^\mu(f)]\Phi) = 0 \qquad \forall \Psi, \Phi \in \mathfrak{H}' \tag{2.4.25}$$

and for all test functions $f \in \mathscr{D}(\mathbb{M})$, where, as in the free case, $\mathfrak{H}' \subset \mathfrak{H}$ is the subspace on which the indefinite form $(\cdot,\cdot)$ is positive semidefinite[2].

The next result tells us that the Gupta-Bleuler strategy comes at an unacceptable price [**FePStro74**].

THEOREM 2.12 (No charged fields). *Assume the common domain $\mathfrak{D} \subset \mathfrak{H}'$ of $j_\mu, F_{\mu\nu}$ and ϕ is dense in $\mathfrak{H}'$ and stable under both fields' action. Let the Gupta-Bleuler condition (2.4.25) be satisfied. Then either there are no charged fields ϕ or*

$$(\Phi, \phi(f)\Psi) = 0 \qquad \forall \Psi, \Phi \in \mathfrak{H}', \tag{2.4.26}$$

ie ϕ vanishes weakly on $\mathfrak{H}'$.

PROOF. We define $B^\nu := \partial_\mu F^{\mu\nu} - j^\nu$ and note that on account of Theorem 2.11, we obtain the limit $\lim_{R\to\infty}[B^0(f_u \otimes f_R), \phi(f)] = q\phi(f)$ because $\lim_{R\to\infty}[\partial_\mu F^{\mu 0}(f_u \otimes f_R), \phi(f)] = 0$ from (2.4.24) and the definition of a charged field in (2.4.22). This entails that

$$\lim_{R\to\infty} (\Phi, [B^0(f_u \otimes f_R), \phi(f)]\Psi) = q(\Phi, \phi(f)\Psi). \tag{2.4.27}$$

But because the lhs vanishes for $\Phi, \Psi \in \mathfrak{D}$ due to the Gupta-Bleuler condition $B^\nu(f)\mathfrak{D} = 0$ and $\mathfrak{D} \subset \mathfrak{H}'$ is dense in $\mathfrak{H}'$, we have either $q = 0$ or $(\Phi, \phi(f)\Psi) = 0$ for $\Phi, \Psi \in \mathfrak{H}'$. □

This tells us that neither the strong nor the weak Gupta-Bleuler form of Maxwell's equations with nontrivial currents seem to be compatible with the axiomatic framework.

[2] $(\Phi, \Phi) \geq 0$ for all $\Phi \in \mathfrak{H}'$ but $(\Phi, \Phi) = 0$ does not imply $\Phi = 0$.

2.4.3. Haag's theorem in QED. In the light of the results discussed in this section, which undoubtedly pertain to QED, even though charged scalar field were used, one may suggest two extreme stances regarding Haag's theorem for QED.

(St1) The source-free Maxwell's equations in the sense of the covariant formalism due to Gupta and Bleuler,

$$(\mathfrak{H}_{\text{phys}}, F^{\mu\nu}(\partial_\mu f)\mathfrak{H}_{\text{phys}}) = 0 \qquad \forall f \in \mathscr{S}(\mathbb{M}) \tag{2.4.28}$$

cohere with Wightman's axiomatic framework in the adapted form. Consequently, Haag's theorem applies fully to QED. No one knows whether and in what sense differential equations for the interacting operator fields are fulfilled. We should therefore not impose such equations and require only (2.4.28) for the quantisation of the free Maxwell field.

(St2) Maxwell's equations with nontrivial electromagnetic currents

$$(\mathfrak{H}_{\text{phys}}, [F^{\mu\nu}(\partial_\mu f) + j^\nu(f)]\mathfrak{H}_{\text{phys}}) = 0 \qquad \forall f \in \mathscr{S}(\mathbb{M}) \tag{2.4.29}$$

are essential for QED and should not be abandoned. These equations are not compatible with Wightman's axioms even in the gauge-adapted version because they forbid charged fields. Haag's theorem does therefore not hold in the strict sense of coinciding vacuum expectation values in QED[3].

We are inclined to a position that embraces the first stance while leaving room for the second. This is to be understood as follows.

Given the arguments presented in Section 1.5 against the existence of the interaction picture in its canonical form, we are in no doubt that the interaction picture cannot exist in QED. The UV divergences one encounters in perturbation theory of QED show what happens if one assumes otherwise. Yet the canonical theory does not stop there but changes the rules of the game drastically by renormalisation.

While unrenormalised QED almost surely falls prey to Haag's theorem, we contend that *renormalised*, ie physical QED, which has to be clearly distinguished from its unrenormalised cousin, is safe from it:

1. Renormalised QED yields nontrivial results in perturbation theory that agree nicely with experiment.
2. As will become clear in Chapter 3, Section 3.5, we contend that Haag's theorem cannot be applied to renormalised QED for the very reason that it is *almost surely not unitary equivalent to a free theory*, where we remind the reader that this form of equivalence is one of Haag's theorem's core provisions.
3. The validity of Wightman's axiomatic framework is dubitable given the results by Strocchi and collaborators discussed in this section, especially Theorem 2.12.

Besides, and this opens flanking fire against one of Wightman's axioms in QED which (to the ken of the author) seems to have been so far been overlooked: the spectral condition for photons (Axiom I). Because the t-channel of fermion interactions contributes to the four-point scattering amplitude in QED, one simply needs and routinely utilises the concept of *virtual, that is, spacelike photons*[4]! If we consider what the spectral condition (Axiom I) entails for the vacuum expectation values of a general QFT and take into account the results that perturbative approaches have brought to light so far, *the contrast could hardly be starker: the spectral condition for photons is probably never satisfied*, at least to the best of our knowledge!

Regarding Maxwell's equations, a possibly existent reconstructed renormalised QED may have observables which in some subspace of $\mathfrak{H}_{\text{phys}}$ satisfy Maxwell's equations, albeit within a form of the Wightman framework that, to this day, is still inconceivable.

[3]Haag's theorem relies heavily on the validity of the Wightman framework.

[4]See also Section (6.6) on spacelike photons.

CHAPTER 3

Renormalisation and Haag's theorem

Haag's theorem directly refutes the *Gell-Mann-Low formula*, a result which was derived in [**GeMLo51**] and became widely known as the central assertion of the *Gell-Mann-Low theorem*. We shall quickly review it in Section 3.1 to see how the arguments used in the proofs of both theorems relate.

Since the Gell-Mann-Low formula purports to relate the time-ordered n-point functions of two intertwined field theories, one may try to tackle the *CCR question* for the interacting field. Section 3.2 shows that the Gell-Mann-Low formula has no answer for the very reason that it requires the time ordering of observables.

The divergences encountered in perturbation theory clearly signify that the Gell-Mann-Low theorem cannot possibly be applicable to unrenormalised QFT. In Section 3.3, we review how these UV divergences of φ^4-theory are incurred and why regularisation alone is not acceptable for a QFT. Even if one chooses a regularisation method that preserves Poincaré invariance can the resulting theory not be unitary equivalent to a free theory, a simple consequence of Haag's theorem.

Because of the somewhat awkward way renormalisation is introduced in canonical QFT, it presents itself more like a narrative than a theory, as narrated in Section 3.4. However, it is the best narrative we have and we shall argue that counterterms can be seen as auxiliary interaction terms that capture the complexity of relativistic interactions at least in some sense and are needed to make up for the 'wrong choice' of Lagrangian.

Finally, in Section 3.5 we present an argument which next-to proves that the intertwiner between the free and the interacting field theory cannot be unitary. Because a unitary intertwiner is essential in Haag's theorem, we conclude that a renormalised (scalar) field theory is not subject to this triviality dictum.

3.1. The theorem of Gell-Mann and Low

The operations performed in the following derivation of the Gell-Mann-Low formula are purely formal and not well-defined in quantum field theory (QFT). Gell-Mann and Low simple assume that both free and interacting Hamiltonians are given as bona fide operators on a common Hilbert space $\mathfrak{H}$ with their individual ground states, ie the vacua.

3.1.1. Review of the interaction picture. We first remind ourselves of the 3 pictures in quantum theory, namely Schrödinger, Heisenberg and interaction picture. The latter is also known as Dirac picture. Let

$$\varphi(t,\mathbf{x}) = e^{iHt}\varphi(\mathbf{x})e^{-iHt} \qquad \text{(Heisenberg picture)} \tag{3.1.1}$$

be the Heisenberg picture field and

$$\varphi_0(t,\mathbf{x}) = e^{iH_0t}\varphi(\mathbf{x})e^{-iH_0t} \qquad \text{(interaction picture)} \tag{3.1.2}$$

the interaction picture field both at time t, where $\varphi(\mathbf{x})$ is the time-independent Schrödinger picture field, H the Hamiltonian of the full interacting theory and H_0 that of the free theory.

Both pictures are consequently intertwined according to

$$\varphi(t,\mathbf{x}) = e^{iHt}\underbrace{e^{-iH_0t}\varphi_0(t,\mathbf{x})e^{iH_0t}}_{\varphi(\mathbf{x})}e^{-iHt} = V(t)^\dagger\varphi_0(t,\mathbf{x})V(t) \tag{3.1.3}$$

where the operator fields coincide at $t = 0$. The idea is borrowed from classical mechanics: from looking at a particle system on a time slice one cannot infer whether its constituents interact. This is only possible by watching how things change in the course of time, ie how the system *evolves* in time. Some authors, especially in older textbooks like [**ItZu80**], replace $V(t)$ in (3.1.3) by the time-ordered exponential

$$\mathsf{U}(t,-\infty) = \mathsf{T}e^{-i\int_{-\infty}^{t} d\tau\ H_I(\tau)} \tag{3.1.4}$$

where $H_I(t)$ is the interacting part of the Hamiltonian in terms of the incoming free field φ_{in} which takes the role of φ_0 in their treatment. This incoming free field then agrees with the Heisenberg field in the remote past $t \to -\infty$ and not, as in our case, on a time slice.

Notwithstanding this detail, both formulations purport to employ a unitary map that relates the interacting Heisenberg field φ to the free interaction picture field φ_0 such that (3.1.3) holds for any time. Notice that Haag's theorem asks for less, namely that the unitary relation is given at one fixed instant.

To recall how the interaction picture states are defined and evolved in time, we consider the expectation value

$$\langle\Psi|\varphi(t,\mathbf{x})\Psi\rangle = \langle\Psi|V(t)^\dagger\varphi_0(t,\mathbf{x})V(t)\Psi\rangle = \langle V(t)\Psi|\varphi_0(t,\mathbf{x})V(t)\Psi\rangle, \tag{3.1.5}$$

where Ψ is a stationary reference Heisenberg state. This expression suggests that $\Psi(t) = V(t)\Psi$ is an interaction picture state at time t, evolved in this picture from Ψ. A transition from one interaction picture state $\Psi(t)$ into another $\Psi(s)$ at time s, is thus governed by the evolution operator $\mathsf{U}(t,s)$ given by

$$\Psi(t) = V(t)\Psi = V(t)V(s)^\dagger\Psi(s) =: \mathsf{U}(t,s)\Psi(s). \tag{3.1.6}$$

The interaction picture state $\Psi(s)$ at time s is thus time-evolved to time t by the operator

$$\mathsf{U}(t,s) = V(t)V(s)^\dagger = e^{iH_0t}e^{-iH(t-s)}e^{-iH_0s}. \tag{3.1.7}$$

Let Ω_0 be the vacuum of H_0, ie $H_0\Omega_0 = 0$, $\{\Psi_n\}$ an eigenbasis of the Hamiltonian H and E_n the corresponding eigenvalues, ie $H\Psi_n = E_n\Psi_n$. The identity operator $\mathrm{id}_{\mathfrak{H}}$ is assumed to have a spectral decomposition which we write as $\mathrm{id}_{\mathfrak{H}} = \sum_{n\geq 0}\mathsf{E}_n$, in which $\mathsf{E}_n = \langle\Psi_n|\cdot\,\rangle\Psi_n$ are the projectors of the presumed energy eigenbasis. $E_0 = 0$ is the ground state energy and Ψ_0 the vacuum of H.

Now here is a crucial identity for the Gell-Mann-Low formula: the two vacua Ψ_0 and Ω_0 are mapped into each other by

$$\mathsf{U}(t,\pm\infty)\Omega_0 = c_0V(t)\Psi_0, \tag{3.1.8}$$

with $c_0 := \langle\Psi_0|\Omega_0\rangle$ being the overlap between the two vacua. This is made plausible by considering the following computation

$$\mathsf{U}(t,s)\Omega_0 = V(t)e^{iHs}\Omega_0 = \sum_{n\geq 0}V(t)e^{iHs}\mathsf{E}_n\Omega_0 = c_0V(t)\Psi_0 + \sum_{n\geq 1}V(t)e^{iE_ns}\langle\Psi_n|\Omega_0\rangle\Psi_n, \tag{3.1.9}$$

where we have used $H_0\Omega_0 = 0$, slipped in the spectral decomposition of the identity operator and utilised $H\Psi_0 = 0$. The limit $s \to \pm\infty$ then forces the remainder of the sum to vanish on account of the Riemann-Lebesgue lemma from complex analysis[1] (by 'analogy' because $E_n \neq 0$ for $n > 0$). Then (3.1.8) follows.

[1]We prefer this argument to the one often used when letting $s \to \infty$ in $s(1 + i\varepsilon)$ in (3.1.9), cf.[**PeSch95**]

3.1.2. Gell-Mann-Low formula. For the two-point function, this entails the following. First consider[2]

$$
\begin{aligned}
\langle\Psi_0|\varphi(x_1)\varphi(x_2)\Psi_0\rangle &= \langle\Psi_0|V(t_1)^\dagger\varphi_0(x_1)V(t_1)V(t_2)^\dagger\varphi_0(x_2)V(t_2)\Psi_0\rangle \\
&= \langle V(t_1)\Psi_0|\varphi_0(x_1)V(t_1)V(t_2)^\dagger\varphi_0(x_2)V(t_2)\Psi_0\rangle \\
&= |c_0|^{-2}\langle \mathsf{U}(t_1,+\infty)\Omega_0|\varphi_0(x_1)\mathsf{U}(t_1,t_2)\varphi_0(x_2)\mathsf{U}(t_2,-\infty)\Omega_0\rangle \\
&= |c_0|^{-2}\langle\Omega_0|\mathsf{U}(+\infty,t_1)\varphi_0(x_1)\mathsf{U}(t_1,t_2)\varphi_0(x_2)\mathsf{U}(t_2,-\infty)\Omega_0\rangle.
\end{aligned}
\tag{3.1.10}
$$

Next, note that no time-ordering is necessary so far and that the constant $c_0 = \langle\Psi_0|\Omega_0\rangle$ can be expressed by using (3.1.8) and, applying the group law $\mathsf{U}(t,s)\mathsf{U}(s,t') = \mathsf{U}(t,t')$, we obtain:

$$
|c_0|^2 = \langle\mathsf{U}(t,+\infty)\Omega_0|\mathsf{U}(t,-\infty)\Omega_0\rangle = \langle\Omega_0|\mathsf{U}(+\infty,-\infty)\Omega_0\rangle = \langle\Omega_0|\mathsf{S}\Omega_0\rangle,
\tag{3.1.11}
$$

where $\mathsf{S} := \mathsf{U}(+\infty,-\infty)$ is the S-matrix in the interaction picture. For the next step we are coerced to time-order the two field operators! Only once this is done can we piece together the S-matrix from the evolution operators in the last line of (3.1.10) to replace it by the *time-ordered product*, denoted by $\mathsf{T}\{...\}$, ie

$$
\begin{aligned}
&\mathsf{U}(+\infty,t_1)\varphi_0(x_1)\mathsf{U}(t_1,t_2)\varphi_0(x_2)\mathsf{U}(t_2,-\infty) \\
&\qquad = \mathsf{T}\{\mathsf{U}(+\infty,t_1)\mathsf{U}(t_1,t_2)\mathsf{U}(t_2,-\infty)\varphi_0(x_1)\varphi_0(x_2)\} \\
&\qquad = \mathsf{T}\{\mathsf{U}(+\infty,-\infty)\varphi_0(x_1)\varphi_0(x_2)\} = \mathsf{T}\{\mathsf{S}\ \varphi_0(x_1)\varphi_0(x_2)\}
\end{aligned}
\tag{3.1.12}
$$

and arrive at

$$
\langle\Psi_0|\mathsf{T}\{\varphi(x_1)\varphi(x_2)\}\Psi_0\rangle = \frac{\langle\Omega_0|\mathsf{T}\{\mathsf{S}\ \varphi_0(x_1)\varphi_0(x_2)\}\Omega_0\rangle}{\langle\Omega_0|\mathsf{S}\Omega_0\rangle}.
\tag{3.1.13}
$$

For the n-point functions this is easily generalised to

$$
\langle\Psi_0|\mathsf{T}\{\varphi(x_1)...\varphi(x_n)\}\Psi_0\rangle = \frac{\langle\Omega_0|\mathsf{T}\{\mathsf{S}\ \varphi_0(x_1)...\varphi_0(x_n)\}\Omega_0\rangle}{\langle\Omega_0|\mathsf{S}\Omega_0\rangle} \quad \text{(Gell-Mann-Low formula)}
\tag{3.1.14}
$$

which finally is the Gell-Mann-Low formula.

Haag's theorem directly controverts this formula or at least says $\mathsf{S}\Omega_0 = \langle\Omega_0|\mathsf{S}\Omega_0\rangle\Omega_0$, ie that provided the above S-matrix really exists, then it must act trivially on the vacuum. As this is not acceptable, something must be wrong. In particular, the constant c_0 should vanish if the van Hove phenomenon occurs. Yet canonical perturbation theory depicts the probabilty of this 'vacuum transition' as a divergent series of divergent integrals. The Feynman rules associate these integrals with vacuum graphs such that

$$
|\langle\Omega_0|\Psi_0\rangle|^2 = |c_0|^2 = \langle\Omega_0|\mathsf{S}\Omega_0\rangle = \exp(\textstyle\sum \texttt{vacuum graphs}).
\tag{3.1.15}
$$

However, standard combinatorial arguments now claim that this problematic exponential is cancelled in (3.1.14) since it also appears and fortunately factors out in the numerator of the rhs. So no matter whether the van Hove phenomenon occurs or not, it is irrelevant for the Gell-Mann-Low formula because 'van Hove cancels out'.

Notice that Haag's theorem does not know anything about which interacting Hamiltonian H we choose and how its interaction part

$$
H_{\text{int}} := H - H_0,
\tag{3.1.16}
$$

[2]The conventions of the axiomatic approach in the exposition of Haag's theorem in [**StreatWi00**], which we have also used, and the Gell-Mann-Low formalism differ slightly: V corresponds to $V(t)^\dagger$ and hence V^{-1} to $V(t)$. We apologise for this notational inconvenience.

let alone its interaction picture representation $H_{\mathrm{I}}(t) = e^{iH_0 t} H_{\mathrm{int}} e^{-iH_0 t}$ is concretely constructed. As we have mentioned in two preceding chapters, Haag's theorem does not point out the ill-definedness of an interaction Hamiltonian like

$$H_I(t) = \frac{g}{4!} \int d^3x \; \varphi_0(t,\mathbf{x})^4 =: \int d^3x \; \mathscr{H}_I(t,\mathbf{x}), \tag{3.1.17}$$

which is a monomial of the interaction picture field $\varphi_0(t,\mathbf{x})$.

Haag's theorem instead makes a very general statement, abstracting from the special form the Hamiltonian in a specific scalar theory. All it says is this: any unitary transformation between a free and another sharp-time Wightman field must be such that all their vacuum expectation values agree.

Now, the Gell-Mann-Low formula (3.1.14) asserts the contrary. The reason it does so is that it builds upon the wrong assumption that the interaction picture exists and that the interaction picture's time evolution operator

$$\mathsf{U}(t,s) = e^{iH_0 t} e^{-iH(t-s)} e^{-iH_0 s} = \mathsf{T} e^{-i\int_s^t d\tau H_I(\tau)} \tag{3.1.18}$$

is well-defined.

3.2. The CCR question

We will now briefly discuss the question whether the above presented canonical form of perturbation theory provides the tools to tackle the CCR question for an interacting field φ. Since the Gell-Mann-Low formula (3.1.14) is designed to attain vacuum expectation values for the interacting Heisenberg picture field φ from those of the free interaction picture field φ_0, we may try and employ it. So the question is: does the field φ satisfy the CCR

$$[\varphi(t,f),\varphi(t,g)] = 0 = [\dot\varphi(t,f),\dot\varphi(t,g)], \qquad [\varphi(t,f),\dot\varphi(t,g)] = i(f,g) \tag{3.2.1}$$

(in the spatially smoothed-out form) for all Schwartz functions f,g in space at some time t? We have chosen $\pi(t,f) = \dot\varphi(t,f)$ for the conjugate momentum field which corresponds to Baumann's choice (see Section 1.6).

Whatever the momentum field's form, we may assume that it involves the time derivative. This suffices to conlcude that the Gell-Mann-Low formula is not apt to answer the CCR question. Nor can it be used to show that φ is local. The reason is simply that *time ordering* is indispensible for the Gell-Mann-Low identity (3.1.14).

First consider the case which can formally be treated, namely the first commutator of the CCR,

$$\begin{aligned} \langle\Psi_0|[\varphi(t,f),\varphi(t,g)]\Psi_0\rangle &= \langle\Psi_0|\mathsf{T}\{[\varphi(t,f),\varphi(t,g)]\}\Psi_0\rangle \\ &= \frac{\langle\Omega_0|\mathsf{T}\{\mathsf{S}\;[\varphi_0(t,f),\varphi_0(t,g)]\}\Omega_0\rangle}{\langle\Omega_0|\mathsf{S}\Omega_0\rangle} = 0, \end{aligned} \tag{3.2.2}$$

because $[\varphi_0(t,f),\varphi_0(t,g)] = 0$ holds for the free field φ_0 and time ordering does not change anything in the first step. If Ψ_0 is cyclic for the field algebra of φ, one may argue that additionally inserting any number of already appropriately time-ordered field operators between the commutator and the two vacua on the lhs of (3.2.2) does not change the fact that the corresponding rhs vanishes. This then entails $[\varphi(t,f),\varphi(t,g)] = 0$, ie φ exhibits a weak form of locality one might call *time-slice locality.*

To tackle the other commutators of the CCR, let us next consider

$$\frac{1}{\varepsilon}[\varphi(t,f),\varphi(t+\varepsilon,g)-\varphi(t,g)] = \frac{1}{\varepsilon}[\varphi(t,f),\varphi(t+\varepsilon,g)]. \tag{3.2.3}$$

Because this expression vanishes when time-ordered, whatever sign $\varepsilon \neq 0$ takes, we cannot apply the Gell-Mann-Low formula as it relies on time ordering. In other words, even if we try weaker concepts of differentiation like left and right derivatives, ie taking the limits '$\varepsilon \uparrow 0$' or

'$\varepsilon \downarrow 0$' instead of '$\varepsilon \to 0$', the time-ordering operator renders all these attempts futile. Thus, the CCR question cannot be answered by the Gell-Mann-Low formula and consequently also not by perturbation theory as we know it today.

3.3. Divergencies of the interaction picture

Because we encounter prolific divergences when the Gell-Mann-Low formula is expanded in perturbation theory, the contradiction between Haag's theorem and the Gell-Mann-Low formua is resolved. Either the interaction picture is well-defined and trivial or must be ill-defined. In Section 1.5 on Fock space it had already dawned on us that the latter is the case, the divergences only confirm it.

Before we review canonical renormalisation and see how it remedies the divergences in the next section, we remind ourselves in this section how they are incurred in the first place.

3.3.1. Divergences. The problem of defining the interaction part of the Hamiltonian in (3.1.17) appears on the agenda as soon as one attempts to put the Gell-Mann-Low identity (3.1.14) to use in perturbation theory. That is, when the perturbative expansion of the S-matrix, namely *Dyson's series*

$$\mathsf{S} = 1 + \sum_{n\geq 1} \frac{(-i)^n}{n!} \int d^4x_1 \, ... \int d^4x_n \; \mathsf{T}\{\mathscr{H}_I(x_1) \, ... \, \mathscr{H}_I(x_n)\} \tag{3.3.1}$$

is employed (in four-dimensional spacetime).

We have already seen in the discussion of Theorem 2.1, ie Wightman's no-go theorem, that quantum fields are too singular to be defined at sharp spacetime points. Yet we have also seen in Section 2.2 that at least their n-point functions can be given a meaning in the sense of distribution theory (Wightman distributions). However, powers of a free field at one spacetime point are still ill-defined because their vacuum expectation values are divergent, eg

$$\langle \Omega_0 | \varphi_0(x)^2 \Omega_0 \rangle = \infty. \tag{3.3.2}$$

The cure for this lies in defining so-called *Wick powers*, given for a free field (!) recursively by $:\varphi_0(x):= \varphi_0(x)$,

$$:\varphi_0(x)^2 := \lim_{y\to x} \{\varphi_0(x)\varphi_0(y) - \langle \Omega_0|\varphi_0(x)\varphi_0(y)\Omega_0\rangle\} \tag{3.3.3}$$

and

$$:\varphi_0(x)^n := \lim_{y\to x} \{:\varphi_0(x)^{n-1}: \varphi_0(y) - (n-1)\langle \Omega_0|\varphi_0(x)\varphi_0(y)\Omega_0\rangle :\varphi_0(x)^{n-2}:\}, \tag{3.3.4}$$

where the limit is to be understood in the weak sense, ie as a sesquilinear form on Hilbert space [**Stro13**]. In Euclidean field theories, Wick powers are defined mutatis mutandis in the obvious way, ie by the replacement $\langle \Omega_0|...\Omega_0\rangle \to \langle ...\rangle_0$.

However, for operator fields, Wick powers are equivalent to the well-known normal-ordered product in terms of annihilators and creators, namely, in terms of negative and positive frequency pieces,

$$:\varphi_0(x_1) \, ... \, \varphi_0(x_n) := \sum_{J\subseteq\{1,...,n\}} \prod_{j\in J} \varphi_0^-(x_j) \prod_{i\notin J} \varphi_0^+(x_i) \tag{3.3.5}$$

where the limit $x_j \to x$ is subsequently taken inside an expectation value.

Products of Wick-ordered monomials evaluate to a product of free two-point functions which are well-defined in the sense of distributions [**BruFK96**]:

$$\langle \Omega_0| :\varphi_0(x_1)^{n_1}: ... :\varphi_0(x_k)^{n_k}: \Omega_0\rangle = \sum_{G\in\mathcal{G}(n_1,...,n_k)} c(G) \prod_{l\in E(G)} \Delta_+(x_{s(l)} - x_{t(l)}; m^2), \tag{3.3.6}$$

in which the notation has the following meaning:

- $\mathcal{G}(n_1, ..., n_k)$ is the set of all directed graphs without self-loops consisting of k vertices with valencies $n_1, ..., n_k$, respectively,

ie the i-th vertex, associated with the spacetime point $x_i \in \mathbb{M}$, has n_i lines attached to it.

- $E(G)$ is the edge set of the graph G,
- $s(l)$ and $t(l)$ are source and target vertex of the line $l \in E(G)$.

The factor $c(G)$ is of purely combinatorial nature and is not of import to our discussion here (see [**BruFK96**]).

Thus, one may alter the definition of the interaction picture Hamiltonian (3.1.17) into the Wick-ordered form

$$\mathscr{H}_I(x) = \frac{g}{4!} : \varphi_0(x)^4 : \tag{3.3.7}$$

to better the understanding of Dyson's series (3.3.1). In terms of Feynman graphs, this means that self-loops evaluate to zero.

Yet the Gell-Mann-Low formula (3.1.14) and Dyson's series (3.3.1) require *time-ordered vacuum expectation values*, ie we need the *time-ordered* versions of the (3.3.6) which by virtue of Wick's theorem [**Wic50**] evaluates formally to [**Fred10**]

$$\langle \Omega_0 | \mathsf{T}\{: \varphi_0(x_1)^{n_1} : ... : \varphi_0(x_k)^{n_k} :\}\Omega_0\rangle = \sum_{G \in \mathcal{G}(n_1,...,n_k)} c(G) \prod_{l \in E(G)} i\Delta_F(x_{s(l)} - x_{t(l)}; m^2) \tag{3.3.8}$$

and requires us to use time-ordered two-point functions, known as Feynman propagators (in position space):

$$i\Delta_F(x-y; m^2) := \langle \Omega_0 | \mathsf{T}\{\varphi_0(x)\varphi_0(y)\}\Omega_0\rangle, \tag{3.3.9}$$

given by the distribution

$$\Delta_F(x-y; m^2) = \lim_{\epsilon \downarrow 0} \int \frac{d^4p}{(2\pi)^4} \frac{e^{-ip(x-y)}}{p^2 - m^2 + i\epsilon}. \tag{3.3.10}$$

Note that the Feynman propagator has the property $\Delta_F(x-y; m^2) = \Delta_F(y-x; m^2)$, on account of the time-ordering. As is well known, products of these objects are in general ill-defined and are the origin of UV divergences in perturbation theory [**He66**], in contrast to products of *Wightman distributions* $\Delta_+(x-y; m^2) = \langle \Omega_0 | \mathsf{T}\{\varphi_0(x)\varphi_0(y)\}\Omega_0\rangle$. Thus, the healing effect of Wick ordering has been reversed by the time-ordering.

If we nevertheless insert Dyson's series (3.3.1) into the Gell-Mann-Low formula (3.1.14) and use the interaction Hamiltonian (3.3.7), we get

$$\begin{aligned} &\langle \Psi_0 | \mathsf{T}\{\varphi(x_1)...\varphi(x_n)\}\Psi_0\rangle \\ &= \frac{1}{|c_0|^2} \sum_{l \geq 0} \frac{(-ig)^l}{(4!)^l l!} \int d^4y_1 \, ... \int d^4y_l \, \langle \Omega_0 | \mathsf{T}\{: \varphi_0(y_1)^4 : \ ... : \varphi_0(y_l)^4 : \ \varphi_0(x_1)...\varphi_0(x_n)\}\Omega_0\rangle \end{aligned} \tag{3.3.11}$$

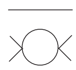
which is a (formal) power series in the parameter g. It is ill-defined even if viewed as an asymptotic series: only its first few coefficients exist while the remainder consists of badly divergent integrals. In view of Haag's theorem, this is no surprise, though. We would, in fact, be confronted with a serious puzzle had we found a well-defined expression! Luckily, (3.3.11) is ill-defined.

3.3.2. Regularisation. Contrary to the commonly adopted view, the *combinatorial approach*, to be expounded carefully in Chapter 4, takes the following pragmatic stance. What (3.3.11) confronts us with, is an expression containing *combinatorial data about a certain class* *of distributions* in the form of a formal power series. In this sense it is not meaningless. Let us simply write the series (3.3.11) as

$$\langle \Psi_0 | \mathsf{T}\{\varphi(x_1)...\varphi(x_n)\}\Psi_0\rangle = \sum_{G\in\mathcal{G}_n} g^{|V(G)|} \prod_{\ell\in L(G)} i\Delta_F(\ell) \prod_{\gamma\in C(G)} (\mathbb{M}_\gamma, \nu_\gamma), \tag{3.3.12}$$

where

- $\mathcal{G}_n$ is the set of all scalar Feynman graphs, disconnected as well as connected, with n external ends and vertices of the four-valent type, ie ' $\times$ '. An example is the graph

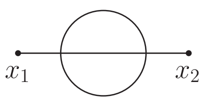
(3.3.13) $G =$ 

  which has $n = 6$ external ends and $|V(G)| = 2$ vertices.
- $L(G)$ is the set of connected pieces with no vertex, ie freely floating lines which connect two external points (the example graph G has one such line).
- $C(G)$ is the set of all connected pieces contained in the graph G with at least one vertex (G has one such piece),
- $V(G)$ is the vertex set of G and $|V(G)|$ its cardinality.

The symbol $i\Delta_F(\ell)$ is a shorthand for the Feynman propagator (3.3.10) associated to the line $\ell \in L(G)$. The pair $(\mathbb{M}_\gamma, \nu_\gamma)$, referred to as *formal pair*, stands for the corresponding divergent integral as follows: the first component $\mathbb{M}_\gamma = \mathbb{M}^{|\gamma|}$ is the domain of integration[3], while the second is the integrand written as a differential form. If the integral is convergent, we identify the formal pair with the integral it represents and write

$$(\mathbb{M}_\gamma, \nu_\gamma) = \int_{\mathbb{M}_\gamma} \nu_\gamma =: \int \nu_\gamma. \tag{3.3.14}$$

For the sake of a neater and parsimonious notation, we suppress the dependence on the spacetime points $x_1, ..., x_n$. In the case $n = 2$ and $v = 2$ we have, for example, the connected Feynman graph

(3.3.15) $\gamma =$ 

with differential form

$$\nu_\gamma(x_1, x_2, y_1, y_2) = -\frac{1}{3!} i\Delta_F(x_1 - y_1)\,(i\Delta_F(y_1 - y_2))^3\, i\Delta_F(y_2 - x_2) d^4y_1 d^4y_2 \tag{3.3.16}$$

and $\mathbb{M}_\gamma = \mathbb{M}^2$ for the two Minkowski integration variables $y_1, y_2 \in \mathbb{M}$. Clearly, the corresponding formal pair is not a convergent integral which cannot be given a meaning as a distribution.

In cases like this where a formal pair represents a divergent integral, one must *regularise* it. This is done in various ways. All regularisation methods have in common that they alter the differential form[4]. Not all of them have a clear physical interpretation. If we take $h_\varepsilon \in \mathscr{D}(\mathbb{M})$, ie a Schwartz function of compact support such that $h_\varepsilon(x) = 0$ for all $x \in \mathbb{M}$ with Euclidean length $||x|| < \varepsilon$, then

$$\Delta_F^\varepsilon(x) := h_\varepsilon(x)\Delta_F(x) \tag{3.3.17}$$

is a nicely behaving regularised Feynman propagator. Products of (3.3.17) can be freely integrated. This non-standard regularisation, which concerns us here only for the sake of the investigation, has eliminated two problems. First, by letting h_ε have compact support, we stave off infrared divergences inflicted by the infinite volume of spacetime $\mathbb{M}$. Second, because h_ε

[3] $|\gamma|$ denotes the loop number of the connected graph $\gamma \in C(G)$.

[4] Dimensional regularisation is no exception.

vanishes on a neighbourhood of the origin, we are also save from ultraviolet (UV) singularities, ie short-distance singularities.

The regularised version of the differential form in (3.3.16), with all Feynman propagators replaced by regularised ones, can now be construed as the distribution

$$f \mapsto \int \nu_\gamma^\varepsilon(f) := \int d^4x_1 \int d^4x_2 \int \nu_\gamma^\varepsilon(x_1, x_2, y_1, y_2) f(x_1, x_2) \tag{3.3.18}$$

for $f \in \mathscr{S}(\mathbb{M}^2)$.

When all formal pairs in (3.3.12) are regularised, now denoted by $(\mathbb{M}_\gamma, \nu_\gamma^\varepsilon) = \int \nu_\gamma^\varepsilon$, one obtains an asymptotic series in the coupling g with coefficients representing distributions,

$$\langle \Psi_0 | \mathsf{T}\{\varphi(x_1)...\varphi(x_n)\} \Psi_0 \rangle_\varepsilon := \sum_{G \in \mathcal{G}_n} g^{|V(G)|} \prod_{\ell \in L(G)} i\Delta_F^\varepsilon(\ell) \prod_{\gamma \in C(G)} \int \nu_\gamma^\varepsilon, \tag{3.3.19}$$

to be applied to a test function $f \in \mathscr{S}(\mathbb{M}^n)$. Once that is done, one arrives at a formal power series with complex numbers as coefficients.

Suppose we had a suitable resummation scheme for this series, then the result may enable us to define the rhs of (3.3.19) as a distribution. But its dependence on the regularisation function h_ε is unacceptable, not least because Poincaré invariance is violated. To get rid of this dependence, the *adiabatic limit* $\lim_{\varepsilon \to 0} h_\varepsilon = 1$ is necessary, a condition that we additionally impose on h_ε. Of course, since this restores the unfavourable original situation of divergent integrals, one has to modify the formal pairs in such a way that their limits lead to convergent and, moreover, Poincaré-invariant integrals.

3.3.3. Evade Haag's theorem by regularisation? To summarise, we note that

- time-ordering, necessitated by the Gell-Mann-Low formula (3.1.14), leads inevitably to ill-defined products of Feynman propagators which then in turn bring about UV divergences;
- although regularisation helps, it is physically unacceptable.

Let us imagine for a moment we had chosen a Poincaré-invariant regularisation method and had found an explanation for why it is physically acceptable and satisfactory. Suppose further that the so-obtained two-point function differs from the two-point function of the free field at spacelike distances. Then, contrary to what some might believe and wish for, by Haag's theorem (Theorem 2.9), we can be sure that the so-reconstructed theory is unitarily inequivalent to the free theory. Invoking the Stone-von Neumann theorem would be futile: something has to give, the provisos of both theorems cannot form a coherent package!

3.4. The renormalisation narrative

We shall review in this section the way renormalisation is nowadays canonically introduced and how it changes the Gell-Mann-Low perturbation expansion so drastically that the formal power series one obtains has finite coefficients. This outcome, however, brings back the conundrum posed by Haag's theorem because the same bold assertions about the interaction picture and the unitarity of its evolution operator are made yet again, albeit this time for the renormalised field.

3.4.1. Counterterms. Now because one cannot accept the regularised theory as the final answer and removing the regulator brings back the divergences, the canonical formalism backpedals at this point. To explain the necessary modifications, *the story is changed* in a decisive way: the coupling g is just the 'bare coupling', employed so far out of ignorance (in a sort of bare state of mind, one might say). The same holds for the *bare mass* m and the *bare field* φ. These 'bare' quantities are deemed unphysical because they have evidently led to divergences.

Dyson explains this situation in [**Dys49b**] by telling the amusing tale of an ideal observer whose measuring apparatus, 'non-atomic' and therefore not comprised of atoms, is only limited by the fundamental constants c and $\hbar$. Performing measurements at spacetime points which the fictitious observer is capable to determine with infinite precision, he finds infinite results.

However, the physical ('renormalised') counterparts of the bare quantities are constructed as follows. First, the bare field gets 'renormalised' by a factor Z:

$$\varphi_r := \frac{\varphi}{\sqrt{Z}}, \tag{3.4.1}$$

where the resulting field φ_r is called *renormalised field*, the new player that takes the place of the 'old', the bare field φ. The so-called *wavefunction renormalisation* or *field-strength renormalisation constant* Z is a function of several variables, in particular of the renormalised coupling g_r. Both renormalised coupling g_r and mass m_r are defined through

$$g = g_r \frac{Z_g}{Z^2}\ , \qquad m = m_r \sqrt{\frac{Z_m}{Z}} \tag{3.4.2}$$

in which two additional renormalisation constants are introduced: Z_g is the *coupling renormalisation constant* and Z_m is the *mass renormalisation constant.* Both are also functions of the renormalised coupling g_r.

We shall now see that when the bare quantities are replaced by their physical, renormalised counterparts, *the net effect is a modified interaction part* of the Hamiltonian and thus of the corresponding Lagrangian. Let us consider the original (bare) Lagrangian, given by

$$\mathscr{L}(\varphi) = \frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}m^2\varphi^2 - \frac{g}{4!}\varphi^4, \tag{3.4.3}$$

formulated in terms of the bare quantities. In terms of the renormalised quantities, this same Lagrangian takes the form

$$\mathscr{L}(\varphi) = \frac{1}{2}Z(\partial\varphi_r)^2 - \frac{1}{2}m_r^2 Z_m\varphi_r^2 - \frac{g_r}{4!}Z_g\varphi_r^4, \tag{3.4.4}$$

where only φ, m and g have been replaced in accordance with (3.4.1) and (3.4.2). The next step is now to split this expression into two pieces: $\mathscr{L}_r$ and what is known as the *counterterm* $\mathscr{L}_{ct}$, that is,

$$\mathscr{L} = \mathscr{L}_r + \mathscr{L}_{ct} \tag{3.4.5}$$

whose components are given by $\mathscr{L}_r = \frac{1}{2}(\partial\varphi_r)^2 - \frac{1}{2}m_r^2\varphi_r^2 - \frac{g_r}{4!}\varphi_r^4$ and the counterterm Lagrangian

$$\mathscr{L}_{ct} = \frac{1}{2}(Z-1)(\partial\varphi_r)^2 - \frac{1}{2}m_r^2(Z_m-1)\varphi_r^2 - \frac{g_r}{4!}(Z_g-1)\varphi_r^4. \tag{3.4.6}$$

The index 'r' in $\mathscr{L}_r$ signifies that this part is composed of the renormalised quantities only. Itzykson and Zuber call $\mathscr{L}$ the 'renormalised Lagrangian' [**ItZu80**]. As this bears some potential for confusion because $\mathscr{L}$ is equal to the original Lagrangian, Itzykson and Zuber admit that this is a very unfortunate denomination (ibidem, p.389).

Unfortunately, this new splitting (3.4.5) of the old Lagrangian marks a handwaving twist in the renormalisation narrative: φ_r is now seen as the proper fully interacting Heisenberg picture field and is subsequently put through the same interaction picture transformation procedure as described in Section 3.1 for the bare field φ. The resulting interaction picture field $\varphi_{r,0}$ is again a free field, this time with mass $m_r \neq m$, but the obvious relation (3.4.1) to the old free field brushed under the carpet. As already alluded to in Section 1.8.2, differing masses imply unitary inequivalence. We shall come back to this point.

The new overhauled narrative starts out with the Lagrangian $\mathscr{L} = \mathscr{L}_{0,r} + \mathscr{L}_{int}$ consisting of new free and interacting parts

$$\mathscr{L} = \underbrace{\frac{1}{2}(\partial\varphi_r)^2 - \frac{1}{2}m_r^2\varphi_r^2}_{\mathscr{L}_{0,r}} \underbrace{-\frac{g_r}{4!}\varphi_r^4 + \mathscr{L}_{ct}}_{\mathscr{L}_{int}}. \tag{3.4.7}$$

The interaction term $\mathscr{L}_{int}$ is now subjected to the same interaction picture procedure as described in Section 3.1 and m_r is the mass of the free interaction picture field. But the crucial difference is that the terms in $\mathscr{L}_{ct}$ have a nontrivial coupling dependence: Z, Z_m and Z_g are themselves seen as functions of the new renormalised coupling g_r and need to be expanded (in perturbation theory).

3.4.2. Renormalised Gell-Mann-Low expansion. As a result, the Gell-Mann-Low expansion is no more an exponential one in the renormalised coupling g_r. At every order of perturbation theory, the counterterm Lagrangian $\mathscr{L}_{ct}$ generates additional divergent integrals (or formal pairs) which, provided the coefficients of the renormalisation constants are chosen correctly, cancel the divergent integrals that the term

$$\mathscr{L}_{int,r} = -\frac{g_r}{4!}\varphi_r^4 \tag{3.4.8}$$

produces (when transformed into the interaction picture representation). The choice of Z, Z_m and Z_g is not unique and needs physical conditions to be fixed[5].

Let the expansions of the renormalisation constants be given by

$$Z(g_r) = 1 + \textstyle\sum_{j\geq 1} a_j g_r^j\ , \qquad Z_m(g_r) = 1 + \sum_{j\geq 1} b_j g_r^j\ , \qquad Z_g(g_r) = 1 + \sum_{j\geq 1} c_j g_r^j. \tag{3.4.9}$$

The condition $Z(0) = Z_m(0) = Z_g(0) = 1$ must be imposed to make sure one obtains the free Lagrangian when setting $g_r = 0$. Let us briefly review the canonical 'song and dance' to see how the coefficients of these series are determined and how they may be interpreted.

One starts by constructing the new renormalised interaction picture Hamiltonian for Dyson's matrix from $\mathscr{L}_{int} = \mathscr{L}_{int,r} + \mathscr{L}_{ct}$ and gets

$$\mathscr{H}_I^r(x) = \frac{g_r}{4!}Z_g(g_r)\varphi_{r,0}(x)^4 + \frac{1}{2}(Z(g_r)-1)(\partial\varphi_{r,0}(x))^2 + \frac{1}{2}m_r^2(Z_m(g_r)-1)\varphi_{r,0}(x)^2 \tag{3.4.10}$$

which is formulated in terms of free interaction picture fields $\varphi_{r,0}$. We may actually interprete this Hamiltonian physically.

The first term describes interactions between particles and cures some of the divergences incurred by vertex corrections, whereas the additional two terms take into account that real-world physical and relativistic interactions change the mass, ie the 'energy-momentum complex' of the system. Technically, their task is to cancel the remainder of the divergences that the first term generates. Yet they do not just cancel divergences, but *bring about a coupling-dependent mass shift*. This almost surely destroys unitary equivalence between the fields φ_r and $\varphi_{r,0}$, as we shall discuss in Section 3.5.

The second step of the canonical procedure is to take the Gell-Mann-Low formula for the renormalised field, ie

$$\langle\Psi_0|\mathsf{T}\{\varphi_r(x_1)...\varphi_r(x_n)\}\Psi_0\rangle = \frac{\langle\Omega_0|\mathsf{T}\{\mathsf{S}_r\ \varphi_{r,0}(x_1)...\varphi_{r,0}(x_n)\}\Omega_0\rangle}{\langle\Omega_0|\mathsf{S}_r\Omega_0\rangle}, \tag{3.4.11}$$

and then trade S_r for Dyson's series

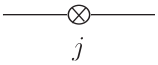
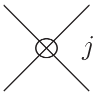
$$\mathsf{S}_r = 1 + \sum_{n\geq 1}\frac{(-i)^n}{n!}\int d^4x_1\ ...\int d^4x_n \mathsf{T}\{\mathscr{H}_I^r(x_1)\ ...\ \mathscr{H}_I^r(x_n)\} \tag{3.4.12}$$

[5]The underlying reason can be illustrated by the fact that a divergent integral $\int I$ is cancelled by $J_C := C - \int I$ for any constant C. It is now a physical choice to specify what C should sensibly be to make sense of $I + J_C$.

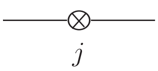
in the renormalised form.

Notice that this series is obtained iteratively in a way that is *independent* of the coupling parameter. Because the coupling dependence of the Hamiltonian has changed now so dramatically, we no longer arrive directly at a perturbative expansion by inserting Dyson's series, as mentioned above: to get there, we have to use the perturbation series of the Z factors in (3.4.9) .

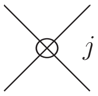
By taking all these perturbative series into account and regularising the resulting Feynman propagators, one arrives at the renormalised analogue of (3.3.19). Its combinatorial content differs substantially, because the renormalised theory has two extra classes of vertex types, associated to the counterterms of the Lagrangian, namely the counterterm vertices

(3.4.13)  and  $j = 1, 2, 3, ...,$

ie an infinite number of different vertices! The number j signifies the power of g_r this graph is associated with. This means in particular that one counterterm vertex of order j counts as if it was a graph with j vertices. These again codify distributions built from integrals,

(3.4.14)  $= i g_r^j \int \frac{d^4p}{(2\pi)^4} [a_j p^2 - m_r^2 b_j] e^{-ip\cdot x}$ (configuration space)

and factors

(3.4.15)  $= -i g_r^j c_j$ (configuration space).

On the grounds that in momentum space, the connected pieces of graphs decompose into 1PI pieces, it is now more convenient, however, to pass over to momentum space by taking the Fourier transform, ie

$$\widetilde{G}_r^{(n)}(p_1, ..., p_n; \varepsilon) := \int dx_1\; e^{ip_1\cdot x_1} ... \int dx_n\; e^{ip_n\cdot x_n} \langle \Psi_0 | \mathsf{T}\{\varphi_r(x_1)...\varphi_r(x_n)\} \Psi_0\rangle_\varepsilon \tag{3.4.16}$$

One obtains

$$\widetilde{G}_r^{(n)}(p_1, ..., p_n; \varepsilon) = (2\pi)^4 \delta^{(4)}(p_1 + ... + p_n) \sum_{G\in\mathcal{G}_n^r} g_r^{|V(G)|} \prod_{\ell\in L(G)} i\widetilde{\Delta}_F^\varepsilon(\ell) \prod_{\gamma\in P(G)} \int \omega_\gamma^\varepsilon, \tag{3.4.17}$$

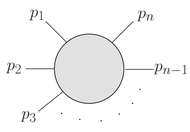
in which

- $\mathcal{G}_n^r$ is the set of all graphs with four-valent vertices and n external legs but this time of the *renormalised* theory, ie including vertices of the class (3.4.13),
- $P(G)$ is the set of all 1PI pieces of the graph G ('P' for proper graphs)

and the set of freely floating lines $L(G)$ is now enriched by external leg (free) propagators.

The new product over all formal pairs $(\mathbb{M}_\gamma, \omega_\gamma^\varepsilon) = \int \omega_\gamma^\varepsilon$ contains the coefficients of the renormalisation Z factors which can be adjusted in such a way that the adiabatic limit $\varepsilon \to 0$ can now be taken without harm in the sense that the result is a formal power series with finite (momentum dependent) coefficients.

The rhs of (3.4.17) is a formal power series with coefficients in the set $\mathscr{S}(\mathbb{M}^n)'$, the set of tempered distributions for which the adiabatic limit does no harm, diagrammatically, we write this as

(3.4.18) $\lim_{\varepsilon\to 0} \widetilde{G}_r^{(n)}(p_1, ..., p_n; \varepsilon) =$  .

Let us assume for simplicity that renormalised φ^4-theory is in some sense Borel-summable[6], ie that the distributions

$$\widetilde{G}^{(n)}_{r,v}(f;\varepsilon) := \sum_{G:|V(G)|=v} \Big(\prod_{\ell\in L(G)} i\Delta^{\varepsilon}_F(\ell) \prod_{\gamma\in C(G)} \int \nu^{\varepsilon}_{\gamma}\Big)(f) \tag{3.4.19}$$

of each vertex level v collectively give rise to formal power series of the form

$$\sum_{v\geq 0} \lim_{\varepsilon\to 0} \widetilde{G}^{(n)}_{r,v}(f;\varepsilon) g_r^v \tag{3.4.20}$$

whose Borel sum

$$\widetilde{G}^{(n)}_r(f,g_r) = \frac{1}{g_r}\int_0^{\infty} e^{-\zeta/g_r} \mathtt{cont}\mathscr{B}\left\{ \textstyle\sum_{v\geq 0} \lim_{\varepsilon\to 0} \widetilde{G}^{(n)}_{r,v}(f;\varepsilon) g_r^v \right\}(\zeta)\; d\zeta \tag{3.4.21}$$

really represents a distribution defined for the test function f. Then, the time-ordered n-point function in configuration space

$$G^{(n)}_r(x_1,...,x_n,g_r) := \langle \Psi_0 | \mathsf{T}\{\varphi_r(x_1)...\varphi_r(x_n)\}\Psi_0\rangle \tag{3.4.22}$$

exists and is given by $G^{(n)}_r(f,g_r) := \widetilde{G}^{(n)}_r(\widetilde{f},g_r)$ for a test function f and its Fourier transform $\widetilde{f}$, both elements in $\mathscr{S}(\mathbb{M}^n)$.

Although this would certainly be a neat result, it brings back the inconvenient question raised by Haag's theorem. According to the canonical narrative, the renormalised free interaction picture field $\varphi_{r,0}$ is unitarily related to the fully interacting renormalised field φ_r by the intertwining relation

$$\varphi_r(t,\mathbf{x}) = \mathsf{T}\{e^{i\int_0^t d^4y\; \mathscr{H}^r_I(y)}\}\varphi_{r,0}(t,\mathbf{x})\mathsf{T}\{e^{-i\int_0^t d^4y\; \mathscr{H}^r_I(y)}\} \tag{3.4.23}$$

in which $\int_0^t d^4y\; \mathscr{H}^r_I(y) := \int_0^t dy_0 \int d^3y\; \mathscr{H}^r_I(y)$. Because the above renormalisation procedure yields finite results, one may argue that this time, the interaction picture has done its job properly and the assertion about unitary equivalence is not contradicted by divergences because there are none.

Of course, there *are* divergences. Notwithstanding that one may argue that the divergent terms cancel each other, Dyson's series (3.4.12) is still not well-defined as the coefficients of the renormalisation factors are divergent integrals themselves. In other words, the formalism is sufficiently dubious such that a finite outcome neither proves nor disproves anything.

3.4.3. Counterterms describe interactions. The renormalised theory with interaction term

$$\mathscr{L}_{int} = -\frac{g_r}{4!}Z_g(g_r)\varphi_r^4 - \frac{1}{2}(Z(g_r)-1)(\partial\varphi_r)^2 - \frac{1}{2}m_r^2(Z_m(g_r)-1)\varphi_r^2. \tag{3.4.24}$$

can therefore not be the final answer. Even though, as explained, the additional counterterms do not seem entirely unphysical, what precludes this Lagrangian description from being a fully satisfactory theory is the fact that the coupling-dependent Z factors in (3.4.24) cannot be chosen finite.

Unfortunately, the hackneyed phrase 'absorption of infinities into couplings and masses' is not just empty but explains nothing physically. Notice that the two mass counterterms by themselves would not produce divergences if their coefficients were finite (see Section 3.5). It is only when they are combined with the vertex interaction term that divergent graphs arise.

As mentioned, these additional terms do more than merely 'counter' and thereby cure the divergences. They had better be seen as some kind of AUXILIARY INTERACTION TERMS WHICH PARTIALLY CAPTURE RELATIVISTIC QUANTUM INTERACTIONS AND COMPENSATE FOR THE ILLS

[6]For the basics of Borel summation and the notation, see Chapter 7 and Appendix Section A.8.

INCURRED BY THE 'WRONG CHOICE' OF LAGRANGIAN. Our motivation for this interpretation is as follows[7].

When relativistic quantum particles interact, they change their mode of existence such that during interactions, the particle concept breaks down completely. Because energy, mass and momentum are intimately related and can only be disentangled for free particles, the initial unrenormalised guess

$$\mathscr{H}_I(x) = \frac{g}{4!}\varphi_0(x)^4 \tag{3.4.25}$$

did not capture the complexity of the relativistic situation. When new particles are created, the mass of the system changes depending on the coupling strength. Consequently, (3.4.25) cannot be sufficient. This, we speculate, may actually be the physical reason behind why Fröhlich has found Euclidean $(\varphi^4)_d$-theories to be trivial for $d \geq 4+\epsilon$ [**Fro82**].

In other words, despite their auxiliary status, we contend that the two additional terms

$$\frac{1}{2}(Z(g_r)-1)(\partial\varphi_{r,0}(x))^2 + \frac{1}{2}m_r^2(Z_m(g_r)-1)\varphi_{r,0}(x)^2 \tag{3.4.26}$$

take into account that relativistic interactions create additional momentum and mass, one at the expensive of the other in a way that depends on the coupling strength.

In a sense, this is an interpretation of the *self-energy* which describes a coupling- and momentum-dependent mass shift. Our interpretation of the counterterms are motivated by the way we think about this very mass shift: to us, it captures relativistic interactions, what else could it possibly say?

Imagine for a moment we had found a Lagrangian not leading to divergences. We would still have to make sure that certain conditions required by physical considerations are satisfied. In the presently known renormalised theories, these take the form of renormalisation conditions. Therefore, we would expect such terms even in a relativistic theory without divergences.

If a mathematically sound Lagrangian quantum field theory is ever possible and one day we succeed in finding the proper interaction term that opens up a viable path to a well-defined perturbative expansion while circumventing and staying clear of divergences all along, we can be sure that the resulting theory is *not* unitary equivalent to a free theory, Haag's theorem is very clear about this.

The underlying reason why the interaction term (3.4.24) has been found is combinatorial in nature. In fact, *there are sound mathematical structures behind renormalisation, namely those of a Hopf algebra.* This was discovered by Kreimer in the late 1990s and further developed in collaboration with Connes and Broadhurst [**Krei02, CoKrei00**].

The feasibilty of the above described renormalisation procedure, proved by Bogoliubov, Parasuik, Hepp and Zimmermann (see [**He66**] and references therein) culminated eventually in what is known as *Zimmermann's forest formula* which solves *Bogoliubov's recursion formula.* Connes and Kreimer later showed that the underlying combinatorics of this latter formula is of Hopf-algebraic nature [**CoKrei00**]. This aspect of renormalisation is explicated in Chapter 4.

3.5. Renormalisation circumvents Haag's theorem

On the assumption that a scalar QFT's n-point distributions can be defined via the limit (3.4.21), one can construe the symbolic expression

$$\varphi_r(t,\mathbf{x}) = \mathsf{T}\{e^{i\int_0^t d^4y\ \mathscr{H}_I^r(y)}\}\varphi_{r,0}(t,\mathbf{x})\mathsf{T}\{e^{-i\int_0^t d^4y\ \mathscr{H}_I^r(y)}\} \tag{3.5.1}$$

[7]An experienced lattice field theorist, whom the author had told about his view, rejected it without giving reasons. He thought it preposterous. But in the absence of such reasons, we find it far from absurd.

as a way to denote the action of the field intertwiner $V_r(t) = \mathsf{T}\{e^{-i\int_0^t d^4y\ \mathscr{H}_I^r(y)}\}$, characterised by the schematic diagram

$$\varphi_{r,0}(x) \quad \xrightarrow{1} \quad \mathsf{S}_r = \mathsf{T}e^{-i\int \mathscr{H}_I^r} \quad \xrightarrow{2} \quad \widetilde{G}_r^{(n)} \xrightarrow{3} \quad \varphi_r(x), \tag{3.5.2}$$

which is to be read as the following procedure:

Step 1: formal construction of Dyson's series S_r from the renormalised interaction picture Hamiltonian $\mathscr{H}_I^r(x)$ of the free field $\varphi_{r,0}(x)$;

Step 2: Gell-Mann-Low expansion, regulator limit at each order of perturbation theory and (some form of) Borel summation which leads to the definition of n-point distributions;

Step 3: reconstruction of the renormalised scalar field theory from the attained n-point distributions by using Wightman's reconstruction theorem.

Because the only provision of Haag's theorem we are not reluctant to give up is *unitary equivalence*, we are inclined strongly to believe that the map $V_r(t)$ cannot be unitary and that this is precisely the reason why

RENORMALISED THEORIES ARE NOT AFFECTED BY HAAG'S TRIVIALITY THEOREM.

The whole canonical narrative just created a misunderstanding by more or less naively nurturing the (certainly not entirely unfounded) belief that one could construct an interacting theory from a free field theory through a unitary intertwining operator $V_r(t)$.

We shall in the following present an argument which makes this unprovable contention *plausible beyond doubt*. The reason it cannot be proved lies in the mathematical ill-definedness of (3.5.1) and that we simply do not know whether the procedure (3.5.2) is feasible.

3.5.1. Mass shift destroys unitary equivalence. We consider a simple toy model given by the Lagrangian

$$\mathscr{L}_m = \frac{1}{2}(\partial\varphi)^2 - \frac{1}{2}m_0^2\varphi^2 - \frac{1}{2}\delta m^2\varphi^2 \tag{3.5.3}$$

with mass shift parameter $\delta m^2 > 0$. This is the example from Duncan's monograph [**Dunc21**] we have mentioned in Subsection 1.8.2. It is clear that the Lagrangian (3.5.3) describes a free field of mass $m > 0$, with $m^2 = m_0^2 + \delta m^2$.

We demonstrate now that if we treat the last term of this Lagrangian as an interaction term and subject it to the canonical procedure of perturbation theory in the interaction picture, the 'mass interaction' term

$$\mathscr{H}_I^m(x) = \frac{1}{2}\delta m^2 : \varphi_0(x)^2 :, \tag{3.5.4}$$

formally obtained by means of the intertwiner

$$V_m(t) = \mathsf{T}\{e^{-i\int_0^t d^4y\ \mathscr{H}_I^m(y)}\} \tag{3.5.5}$$

will then enable us to compute the two-point function through the Gell-Mann-Low expansion

$$\langle\Omega|\mathsf{T}\{\varphi(x)\varphi(y)\}\Omega\rangle = \frac{1}{|c_0|^2}\sum_{n\geq 0}\frac{(-i)^n}{n!}\int_{\mathbb{M}^n} dz\ \langle\Omega_0|\mathsf{T}\{\varphi_0(x)\varphi_0(y)\mathscr{H}_I^m(z_1)...\mathscr{H}_I^m(z_n)\Omega_0\rangle. \tag{3.5.6}$$

We shall now see that the unitary-looking intertwining operator (3.5.5) is not unitary even though it does not at all lead to divergencies!

This then makes it almost certain that renormalised φ^4-theory described by (3.5.2) is unitarily inequivalent to its corresponding free interaction picture field $\varphi_{r,0}$. Since unitary equivalence is a key provision of Haag's theorem, the conclusion is that as a consequence, this theory is hence not afflicted by Haag's theorem.

First, we consider a theorem which we dub 'Haag's theorem for free fields'. This theorem can be found in [**ReSi75**], p.233 (Theorem X.46) . We have simplified it a bit by omitting the conjugate momentum field from the description which makes for a somewhat more straightforward exposition. Note that the dimension of spacetime is of no relevance.

THEOREM 3.1 (Haag's theorem for free fields). *Let φ and φ_0 be two free fields of masses m and m_0, respectively. If at time t there is a unitary map V such that*

$$\varphi(t,\mathbf{x}) = V\varphi_0(t,\mathbf{x})V^{-1}, \tag{3.5.7}$$

then $m = m_0$, ie if $m \neq m_0$ then there exists no such unitary map.

PROOF. The proof is a simple version of the proof of Haag's theorem, Theorem 2.9 (see the arguments there). The conclusion

$$\Delta_+(0,\mathbf{x}-\mathbf{y};m^2) = \Delta_+(0,\mathbf{x}-\mathbf{y};m_0^2) \tag{3.5.8}$$

shows that the assertion is correct and needs no further justification because both fields are free. □

We will now see that canonical perturbation theory enables us to compute the Feynman propagator of the field φ with mass m from the Feynman propagator of φ_0 with mass m_0.

CLAIM 3.2. *Let the symbol $V_m(t) = \mathsf{T}\{e^{-i\int_0^t d^4y\ \mathscr{H}_I^m(y)}\}$ represent the map between the two free fields φ and φ_0 of masses m and m_0, respectively, ie formally,*

$$\varphi(t,\mathbf{x}) = \mathsf{T}\{e^{i\int_0^t d^4y\ \mathscr{H}_I^m(y)}\}\varphi_0(t,\mathbf{x})\mathsf{T}\{e^{-i\int_0^t d^4y\ \mathscr{H}_I^m(y)}\}. \tag{3.5.9}$$

Then the Gell-Mann-Low expansion (3.5.6) yields

$$\langle\Omega|\mathsf{T}\{\varphi(x)\varphi(y)\}\Omega\rangle = \frac{\langle\Omega_0|\mathsf{T}\{\mathsf{S}_m\varphi_0(x)\varphi_0(y)\}\Omega_0\rangle}{\langle\Omega_0|\mathsf{S}_m\Omega_0\rangle} = i\Delta_F(x-y;m^2), \tag{3.5.10}$$

and the map represented by the symbol $V_m(t)$ is not unitary.

PROOF. We compute the rhs of (3.5.6) by using Wick's theorem and obtain

$$\langle\Omega|\mathsf{T}\{\varphi(x)\varphi(y)\}\Omega\rangle = i\Delta_F(x-y;m_0^2) + i\sum_{n\geq 1}(\delta m^2)^n\Delta_F^{*n+1}(x-y;m_0^2) \tag{3.5.11}$$

in which $\Delta_F^{*n+1}(x-y;m_0^2)$ is the $(n+1)$-fold convolution of the Feynman propagator Δ_F given by

$$\begin{aligned}\Delta_F^{*n+1}(x-y;m_0^2) = \int_{\mathbb{M}} dz_1 \ldots \int_{\mathbb{M}} dz_n\ &\Delta_F(x-z_1;m_0^2)\Delta_F(z_1-z_2;m_0^2)\\ &\ldots\ \Delta_F(z_{n-1}-z_n;m_0^2)\Delta_F(z_n-y;m_0^2),\end{aligned} \tag{3.5.12}$$

for $n \geq 1$. The corresponding Feynman diagrams are all of the form

(3.5.13) $G =$ 

.

In momentum space, (3.5.11) takes the simple form

$$\begin{aligned}\int \frac{d^4p}{(2\pi)^4}e^{-ip\cdot x}\langle\Omega|\mathsf{T}\{\varphi(x)\varphi(y)\}\Omega\rangle &= iD_0(p)^{-1} + i\sum_{n\geq 1}(\delta m^2)^n D_0(p)^{-(n+1)}\\ &= \frac{iD_0(p)^{-1}}{1-\delta m^2 D_0(p)^{-1}} = \frac{i}{D_0(p)-\delta m^2} = \frac{i}{D(p)}\end{aligned} \tag{3.5.14}$$

in which $D_0(p) = p^2 - m_0^2 + i0^+$ and $D(p) = p^2 - m^2 + i0^+$ are the inverse free propagators with masses m_0 and m, respectively. By Theorem 3.1, the field interwtiner is not unitary. □

We conclude that the intertwiner between the two fields, symbolically represented by (3.5.5), does indeed transform the field φ_0 into φ but cannot be unitary on account of Theorem 3.1 although the Hamiltonian (3.5.4) is actually a well-defined operator-valued distribution when smeared in space and time. In fact, it is self-adjoint because Wick powers of free fields are [**BruFK96**].

However, the map (3.5.5) is far from being a simple exponentiation of the Hamiltonian $\mathscr{H}_I^m(x)$ because of the integration and time-ordering. It is therefore not surprising that $V_m(t)$ is not unitary.

3.5.2. Counterterms induce mass shift. Because the intertwiner of renormalised φ^4-theory, symbolically given by

$$V_r(t) = \mathsf{T}\{e^{-i\int_0^t d^4y \; \mathscr{H}_I^r(y)}\} \tag{3.5.15}$$

with

$$\mathscr{H}_I^r(x) = \frac{g_r}{4!} Z_g(g_r)\varphi_{r,0}(x)^4 + \frac{1}{2}(Z(g_r)-1)(\partial\varphi_{r,0}(x))^2 + \frac{1}{2}m_r^2(Z_m(g_r)-1)\varphi_{r,0}(x)^2 \tag{3.5.16}$$

also contains a mass shift interaction piece given by the last (two) term(s), we expect it to be also nonunitary. This means that *renormalised φ^4-theory is not affected by Haag's theorem* because the central proviso of unitary equivalence is not fulfilled. Therefore, *renormalisation circumvents this theorem.* If renormalised φ^4-theory exists (which we believe), then it is not difficult to see that the renormalised fully interacting field is certainly not unitary equivalent to a free field by simply taking a look at the mass dependence of its inverse propagator

$$D_r(g_r, p, m_r) = p^2 - m_r^2 - \Sigma_r(g_r, p, m_r) + i0^+ \tag{3.5.17}$$

with self-energy $\Sigma_r(g_r, p, m_r)$. It is this very function which reflects the complexity of relativistic interactions, in some sense correctly captured by the counterterms in (3.5.16).

Given that even the slightest mass shift performed on a free field amounts to a nonunitary transformation to another free field, it is next to impossible for a theory with an interaction dependent (!) mass

$$M^2 = m_r^2 + \Sigma_r(g_r, p, m_r) \tag{3.5.18}$$

to be unitary equivalent to a free field in Fock space of mass m_r.

To assume that an interacting field theory is unitary equivalent is an erroneous idea suggested by the form of the formal identity

$$\varphi(t,\mathbf{x}) = e^{iHt}e^{-iH_0t}\varphi_0(t,\mathbf{x})e^{iH_0t}e^{-iHt} = V(t)^\dagger\varphi_0(t,\mathbf{x})V(t) \tag{3.5.19}$$

for the intertwiner $V(t)$. The canonical formalism works with this assumption in both the renormalised and the unrenormalised case. Only in the unrenormalised case did the fallacy become apparent by emerging divergences.

The renormalised case is different because renormalisation effectively introduces new auxiliary interaction terms so that the theory is changed drastically: by additional (auxiliary) interaction terms called counterterms, a renormalised theory is no longer the unrenormalised renormalisable theory is was prior to renormalisation.

CHAPTER 4

Combinatorial approach: Hopf-algebraic renormalisation

Before the 1950s, renormalisation must have seemed more like a game of whack-a-mole. By 1950, however, it became clear that the rules of renormalisation are not as arbitrary and messy as they appeared to be when they were first formulated [**Dys49a, Dys49b**]. Yet even Feynman, one of the pioneers, was not convinced that renormalisation would be the final answer [**Feyn06**]. This somewhat unsatisfactory situation changed when the concept of the *renormalisation group* was introduced into quantum field theory which offered a new way of looking at the issue [**GeMLo54, Wil75**]. Another perspective, discovered much later by Kreimer and collaborators, unravelled the algebraic underpinnings of renormalisation: it turned out that the rules of the game, if viewed combinatorically in terms of Feynman diagrams, follow the algebraic laws of a *Hopf algebra* [**Krei02, CoKrei98**].

This chapter is a review and at the same time a pedagogical attempt to expound the Hopf-algebraic rules underlying renormalisation without assuming any prior exposure to the algebraic concepts on the part of the reader. Section 4.1 introduces the Hopf algebra of Feynman graphs by means of presenting examples from a scalar theory and QED. Although readable even for readers without any foreknowledge on the algebraic structures, we recommend digesting Appendix Section A.2 first or at least in parallel, as it makes sense to get some familiarity with the Hopf algebra of polynomials in one variable.

This simple yet nontrivial Hopf algebra serves as the paradigm of Appendix A.2. If directly translated to Feynman diagrams, this example corresponds to a situation in which one uses only one divergent primitive graph as a single generator of a Hopf algebra.

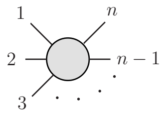
Unlike the way in which Feynman diagrams have been used by physicists before, the combinatorial viewpoint differs profoundly: it treats Feynman diagrams as proper algebraic objects, a 1PI Feynman diagram is identified there with a polynomial variable rather than with an integral [**Krei02**]. As explained in Section 4.2, the connection to Feynman integrals is then mediated in a second step by morphisms known as *Hopf algebra characters*. The associated Appendix Sections A.3 to A.5 have some details and proofs for the assertions stated in the exposition.

We mention for completeness that a vast body of knowledge is available on the Hopf algebra of (decorated) rooted trees, especially regarding Dyson-Schwinger equations [**Foi10, BerKrei06**]. In fact, all of the material in this chapter could equally well be formulated in a language employing only decorated rooted trees. However, to keep this chapter within reasonable bounds, we have decided not to include any of this material as we can do without it. Readers interested in these issues and their applications in QFT are referred to [**CoKrei98, Krei99**].

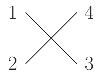
Hopf algebras of Feynman diagrams in general exhibit nontrivial *Hopf ideals* which in the case of gauge theories like QED and QCD correspond to identities known as *Ward-Takahashi* or *Slavnov-Taylor identities*, respectively. This feature is discussed and explained in Section 4.3. It is, however, advisable for readers unacquainted with ideals to first digest Appendix Section A.6 which introduces the concept of a Hopf ideal gently, again through the Hopf algebra of polynomials in one variable. It is shown there that a simple nontrivial Hopf ideal is given by all polynomials with a zero at the origin.

4.1. The Hopf algebra of Feynman graphs

Renormalised perturbation series. We come back to the renormalised perturbation series of φ^4-theory in momentum space,

(4.1.1)
$$(\quad$$

$$)(f) = \sum_{G \in \mathcal{G}_n^r} g_r^{|V(G)|} (\delta_\Sigma \prod_{\ell \in L(G)} i\widetilde{\Delta}_F(\ell) \prod_{\gamma \in P(G)} \int \omega_\gamma)(f),$$

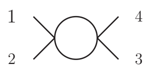
to be understood as the limit (3.4.18), of which we hope that it yields a formal power series representation of a tempered distribution applied to a test function $f \in \mathscr{S}(\mathbb{M}^n)$, as discussed in Section 3.4. The symbol δ_Σ stands for the overall momentum conservation delta distribution. For example, at tree level, we find the distribution

(4.1.2)
$$(\quad$$

$$)(f) = -ig_r \int \prod_{j=1}^{4} \frac{d^4 p_j}{(2\pi)^4} \; \delta_\Sigma(p) \{ \prod_{j=1}^{4} \frac{i}{p_j^2 - m_r^2 + i0^+} \} f(p_1, p_2, p_3, p_4),$$

where $P(G)$ contains in this case only a single vertex: $\gamma = \times$ and $\int \omega_\gamma = -i$. The set $L(G)$ consists of the four external propagators and

$$\delta_\Sigma(p) = (2\pi)^4 \delta^{(4)}(p_1 + p_2 + p_3 + p_4) \tag{4.1.3}$$

is an obvious abbreviation. A less trivial example is

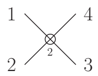
(4.1.4)
$$(\quad$$

$$)(f) = \frac{(-ig_r)^2}{2} \int \prod_{j=1}^{4} \frac{d^4 p_j}{(2\pi)^4} \; \delta_\Sigma(p) \{ \prod_{j=1}^{4} i\widetilde{\Delta}_F(p_j) \} \int \omega (p_1 + p_2) f(p_1, p_2, p_3, p_4),$$

in which

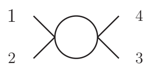
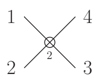
$$\int \omega (p) = \int \frac{d^4 k}{(2\pi)^4} i\widetilde{\Delta}_F(k + p) i\widetilde{\Delta}_F(k) = \int \frac{d^4 k}{(2\pi)^4} \frac{i}{(k+p)^2 - m_r^2 + i0^+} \frac{i}{k^2 - m_r^2 + i0^+} \tag{4.1.5}$$

is the unregularised distribution associated with the graph $\gamma =$. As this section is devoted to the combinatorics of renormalsation, regularisation schemes do not interest us here. Let us for convenience of notation think of all divergent integrals as regularised.

We choose the momentum subtraction (MOM) scheme for renormalisation. The counterterm for (4.1.5) is given by

(4.1.6)
$$(\quad$$

$$)(f) = \frac{c_2}{2} (-ig_r)^2 \int \prod_{j=1}^{4} \frac{d^4 p_j}{(2\pi)^4} \; \delta_\Sigma(p) \{ \prod_{j=1}^{4} i\widetilde{\Delta}_F(p_j) \} f(p_1, p_2, p_3, p_4),$$

where $c_2 = -\int \omega (p_0)$ at some reference momentum $p_0 \in \mathbb{M}$. Then,

(4.1.7)
$$(\quad$$

$$+$$

$$)(f)$$
$$= \frac{(-ig_r)^2}{2} \int \prod_{j=1}^{4} \frac{d^4 p_j}{(2\pi)^4} \; \delta_\Sigma(p) \{ \prod_{j=1}^{4} i\widetilde{\Delta}_F(p_j) \} \int [\omega (p_1 + p_2) - \omega (p_0)] f(p_1, p_2, p_3, p_4),$$

is finite. The differential form $\mathscr{R}[\omega] := \omega (p_1 + p_2) - \omega (p_0)$ has a convergent integrand and is the renormalised differential form for the graph $\gamma =$.

In general, the renormalised differential form $\mathscr{R}[\omega_\gamma]$ of a graph γ is a linear combination of differential forms: we may write it as

$$\mathscr{R}[\omega_\gamma] = \sum_{(\gamma)} \mathscr{C}(\omega_{\gamma'}) \otimes \omega_{\gamma''} \tag{4.1.8}$$

which is one of the central formulae of the Hopf-algebraic approach; when integrated, one gets

$$\int \mathscr{R}[\omega_\gamma] = \int \sum_{(\gamma)} \mathscr{C}(\omega_{\gamma'}) \otimes \omega_{\gamma''} = \sum_{(\gamma)} \int \mathscr{C}(\omega_{\gamma'}) \int \omega_{\gamma''}, \tag{4.1.9}$$

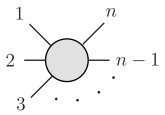
which is what is meant by the tensor product in (4.1.8). The labels γ' and γ'' are elements in the Hopf algebra $\mathcal{H}$ of graphs and $\mathscr{C}$ is a map, to be expounded in due course.

Note that (4.1.8) is essentially what became known as *Bogoliubov's recursion formula* which is solved by *Zimmermann's forest formula*, explained for example in [**ItZu80**]. Suffice it for the moment to acknowledge that in our simple example, we have only two terms in this sum,

$$\mathscr{R}[\omega_{\propto}] = \mathscr{C}(\omega_{\mathbb{I}}) \otimes \omega_{\propto} + \mathscr{C}(\omega_{\propto}) \otimes \omega_{\mathbb{I}}, \tag{4.1.10}$$

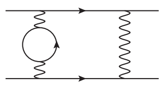
where $\mathscr{C}(\omega_{\mathbb{I}}) = \omega_{\mathbb{I}}$ and $\int \omega_{\mathbb{I}} = 1$, but $\mathscr{C}(\omega_{\propto}) = -\omega_{\propto}|_0$, the latter denoting the evaluation of the integrand at the reference momentum p_0. We see here that the map $\mathscr{C}$ delivers the counterterm. The symbol $\mathbb{I}$ stands for what we shall call the *empty graph*.

The point here is that all data about the renormalised Feynman integral $\int \mathscr{R}[\omega_\gamma]$ is already contained in the set of Feynman diagrams and its Hopf algebra structure of divergent subgraphs. As a consequence, the extra counterterm vertices can be discarded, ie we can replace the set $\mathcal{G}_n^r$ by $\mathcal{G}_n$ in the renormalised perturbation series (4.1.1) and recast it in the form

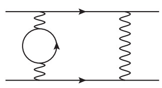
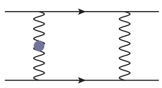
(4.1.11) $\widetilde{G}_r^{(n)}(f) = ($ 

$)(f) = \sum_{G\in\mathcal{G}_n} g_r^{|V(G)|}(\delta_\Sigma \prod_{\ell\in L(G)} i\widetilde{\Delta}_F(\ell) \prod_{\gamma\in P(G)} \int \mathscr{R}[\omega_\gamma])(f).$

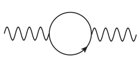
Before we introduce the *Hopf algebra of Feynman graphs* and the algebraic structures underlying renormalisation, let us have a look at an example from quantum electrodynamics (QED): consider the 1PI graph

(4.1.12) $\Gamma =$ 

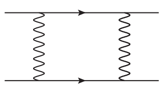
without external propagators. The traditional (diagrammatic) way of writing the renormalised graph $\mathscr{R}(\Gamma)$ is

(4.1.13) $\mathscr{R}(\Gamma) =$  $+$  ,

in which the counterterm 2-vertex cancels the subdivergence

(4.1.14) $\gamma =$  .

If we shrink this subgraph inside the graph Γ (4.1.12) we get the so-called *cograph* of Γ (with respect to γ), denoted by

(4.1.15) $\Gamma/\gamma =$  .

We may write the renormalised differential form for Γ as $\mathscr{R}[\omega_\Gamma] = \omega_\Gamma + c_2\ \omega_{\Gamma/\gamma}$, where the (infinite) constant c_2 must cure the subdivergence of Γ. In the algebraic formulation, this reads

$$\mathscr{R}[\omega_\Gamma] = \mathscr{C}(\omega_{\mathbb{I}}) \otimes \omega_\Gamma + \mathscr{C}(\omega_\Gamma) \otimes \omega_{\mathbb{I}} + \mathscr{C}(\omega_\gamma) \otimes \omega_{\Gamma/\gamma}. \tag{4.1.16}$$

where $\mathscr{C}(\omega_\Gamma) = 0$. This is always the case if Γ is not overall divergent (we will consider this aspect at the end of this section). Again, the map $\mathscr{C}$ takes care of the (sub)divergences and provides the counterterms. We will see that its recursive definition, to be discussed in Section 4.2, reflects the complexity and intricasy of the subdivergence structure of a graph.

4.1.1. Algebra of Feynman graphs. Before we continue and gradually introduce the necessary algebraic structures on the set of Feynman graphs, let us make precise what we mean by a Feynman graph. We will at times also use the term 'Feynman diagram' as a synonym, as we already have.

DEFINITION 4.1 (Feynman graph). *A Feynman graph is a quadruple $G = (V, H, E, \iota)$ which consists of the following data.*

The two basic sets are the vertex set $V(G)$ and the set of half-edges $H(G)$, whereas the set of edges is given as a subset $E(G) \subseteq H(G) \times H(G)$. Finally, the map $\iota\colon H(G) \to V(G)$ tells us which vertex a half-edge is attached to, ie $\iota(h) = v$ if $h \in H(G)$ is anchored in $v \in V(G)$. Furthermore, we call it one-particle irreducible*, or 1PI, if it is connected and deleting an edge leaves it connected.*

We denote the set of all Feynman graphs by $\mathcal{G}$ and split the half-edge set of $G \in \mathcal{G}$ into the set of external and internal half-edges, $H_{ex}(G)$ and $H_{in}(G)$, respectively. Note that this includes all Feynman graphs, both connected and disconnected. Since all we need to construct the perturbation series in (4.1.1) are the 1PI (one-particle irreducible) pieces of a Feynman graph, we focus on these and first generate a commutative algebra

$$\mathcal{H} := \langle\, G \in \mathcal{G} : G \text{ one-particle irreducible } \rangle_{\mathbb{Q}}, \tag{4.1.17}$$

in which the associative and commutative product of two or more graphs is given by the disjoint union, ie

$$\prod_{j\in I} G_j := \bigcup_{j\in I} G_j, \tag{4.1.18}$$

where $I \subset \mathbb{N}$ is a finite index set and $G_j \in \mathcal{G}$ for all $j \in I$. The neutral element of this operation is the empty graph $\mathbb{I} \in \mathcal{H}$, ie $\mathbb{I}G = G\mathbb{I} = G$ for all $G \in \mathcal{H}$.

4.1.2. Coproduct. We now introduce the map that represents the underlying structure for the sum in Bogoliubov's recursion formula (4.1.8), namely the coproduct $\Delta\colon \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$. It takes a graph into a finite sum of elements in $\mathcal{H} \otimes \mathcal{H}$, which we write as

$$\Delta(G) = \sum_{(G)} G' \otimes G'' \tag{4.1.19}$$

for a 1PI Feynman graph G, ie one of the generators of $\mathcal{H}$. By definition, we let it be linear and multiplactive, ie

$$\Delta(\sum_{j\in I} G_j) := \sum_{j\in I} \Delta(G_j), \qquad \Delta(\prod_{j\in I} G_j) := \prod_{j\in I} \Delta(G_j), \tag{4.1.20}$$

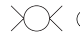
for $G_j \in \mathcal{H}$, $j \in I$, where we recall[1] that $(a \otimes b)(c \otimes d) = ac \otimes bd$ is the product in $\mathcal{H} \otimes \mathcal{H}$. This means

$$\begin{aligned} \Delta(G_1G_2) &= \Delta(G_1)\Delta(G_2) = (\sum_{(G_1)} G_1' \otimes G_1'')(\sum_{(G_2)} G_2' \otimes G_2'') \\ &= \sum_{(G_1)}\sum_{(G_2)} (G_1' \otimes G_1'')(G_2' \otimes G_2'') = \sum_{(G_1)}\sum_{(G_2)} G_1'G_2' \otimes G_1''G_2'' \end{aligned} \tag{4.1.21}$$

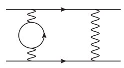
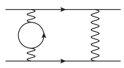
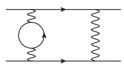
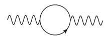
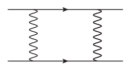
[1]Readers unfamiliar with the tensor product of two linear spaces (or algebras for that matter) are advised to consult Appendix Section A.2.

for the product of two graphs G_1, G_2.

Before we define the coproduct properly, let us take a look at the two examples (4.1.10) and (4.1.16). For the scalar graph, the coproduct yields

(4.1.22) $\Delta($  $) = \mathbb{I} \otimes$  $+$  $\otimes \mathbb{I}$

which corresponds to $\mathscr{R}[\omega] = \mathscr{C}(\omega_{\mathbb{I}}) \otimes \omega + \mathscr{C}(\omega) \otimes \omega_{\mathbb{I}}$ and

(4.1.23) $\Delta($  $) = \mathbb{I} \otimes$  $+$  $\otimes \mathbb{I} +$  $\otimes$ 

reflecting directly $\mathscr{R}[\omega_\Gamma] = \mathscr{C}(\omega_{\mathbb{I}}) \otimes \omega_\Gamma + \mathscr{C}(\omega_\Gamma) \otimes \omega_{\mathbb{I}} + \mathscr{C}(\omega_\gamma) \otimes \omega_{\Gamma/\gamma}$ for the QED example.

Notice what the coproduct does to a graph G: apart from the first two terms $\mathbb{I} \otimes G + G \otimes \mathbb{I}$, which are always produced by the coproduct, it identifies a divergent proper subgraph, puts it in front of the tensor sign and places the corresponding cograph on the right.

We have (not unduly) assumed that the reader knows what a divergent graph is. Let us define it now to introduce the necessary notation. Whether or not a (sub)graph is divergent, depends on what is known as superficial power counting, represented by the function

$$D(G) := d\ |G| + \sum_{e \in E(G)} w(e) + \sum_{v \in V(G)} w(v), \tag{4.1.24}$$

where the map $w : V(G) \cup E(G) \to \mathbb{R}$ gives the *weight* of the edges and vertices and depends on the theory in question (eg for scalar graphs we have $w(e) = -2$ and $w(v) = 0$). d is the spacetime dimension. Here is the definition of a divergent graph.

DEFINITION 4.2. *A 1PI graph $G \in \mathcal{G}$ is called (superficially) divergent if $D(G) \geq 0$. We say that it is logarithmically divergent if $D(G) = 0$, linearly divergent if $D(G) = 1$ and quadratically divergent in case $D(G) = 2$.*

We shall for convenience drop the term 'superficial', as we have always done before. Let

$$Q(G) := \{\gamma \subsetneq G : \gamma = \Pi_j \gamma_j,\ \gamma_j \text{ 1PI and } D(\gamma_j) \geq 0\ \forall j\} \tag{4.1.25}$$

be the set of all proper subgraphs which are a product of divergent 1PI subgraphs. We are now ready to define the coproduct of the 1PI graph $\Gamma \in \mathcal{G}$ by

$$\Delta(\Gamma) := \mathbb{I} \otimes \Gamma + \Gamma \otimes \mathbb{I} + \sum_{\gamma \in Q(\Gamma)} \gamma \otimes \Gamma/\gamma, \tag{4.1.26}$$

where Γ/γ is the cograph, ie the graph one arrives at upon shrinking of γ in Γ to a point which then forms a single vertex in case γ is not a propagator graph. If γ is a propagator graph, it is simply replaced by an internal line. To be combinatorically more precise, in terms of the half-edge and vertex sets, we have

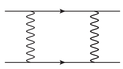
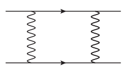
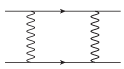
$$H(\Gamma/\gamma) = \begin{cases} H(\Gamma) - H(\gamma)\ , & \gamma \text{ propagator graph, ie } |H_{ex}(\gamma)| = 2 \\ H(\Gamma) - H_{in}(\gamma)\ , & \gamma \text{ vertex graph, ie } |H_{ex}(\gamma)| > 2 \end{cases}, \tag{4.1.27}$$

and

$$V(\Gamma/\gamma) = \begin{cases} V(\Gamma) - V(\gamma)\ , & \gamma \text{ propagator graph, ie } |H_{ex}| = 2 \\ [V(\Gamma) - V(\gamma)] \cup \{v_\gamma\}\ , & \gamma \text{ vertex graph, ie } |H_{ex}| > 2 \end{cases} \tag{4.1.28}$$

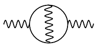
where v_γ is the new vertex replacing the vertex graph γ in Γ.

Hopf algebra elements $p \in \mathcal{H}$ with a coproduct of the form such that $\Delta(p) = \mathbb{I} \otimes p + p \otimes \mathbb{I}$ are referred to as *primitive*. An example is

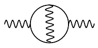
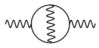
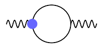
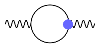
(4.1.29) $\Delta($  $) = \mathbb{I} \otimes$  $+$  $\otimes \mathbb{I}.$

Another example of a primitive graph is $p =$, as we have seen in (4.1.22). This is not the most simple example though. Yet simpler is $\Delta(\mathbb{I}) = \mathbb{I} \otimes \mathbb{I}$, whereas $\Delta(0) = 0$ (by linearity) cannot be undercut in simplicity. Elements $g \in \mathcal{H}$ such that $\Delta(g) = g \otimes g$ are called *grouplike*. The null element is both grouplike and primitive.

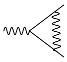
The case of overlapping divergences is treated as follows. Consider the QED 2-loop graph

(4.1.30) $\Gamma =$  .

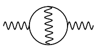
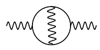
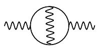
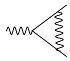
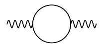
To renormalize it, one needs three counterterms:

(4.1.31) $\mathscr{R}($  $) =$  $+$  $+$  $+$,

where the second and third terms remedy the vertex subdivergence(s) given by the subgraph

(4.1.32) $\gamma =$ 

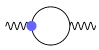
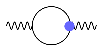
and the last term deals with the overall divergence. In terms of the coproduct, this takes the form

(4.1.33) $\Delta($  $) = \mathbb{I} \otimes$  $+$  $\otimes \mathbb{I} + 2$  $\otimes$  .

and in terms of differential forms,

$$\mathscr{R}[\omega_\Gamma] = \mathscr{C}[\omega_{\mathbb{I}}] \otimes \omega_\Gamma + \mathscr{C}[\omega_\Gamma] \otimes \omega_{\mathbb{I}} + 2\, \mathscr{C}[\omega_\gamma] \otimes \omega_{\Gamma/\gamma}, \tag{4.1.34}$$

where

(4.1.35) $2 \int (\mathscr{C}[\omega_\gamma] \otimes \omega_{\Gamma/\gamma}) = 2 \int \mathscr{C}[\omega_\gamma] \int \omega_{\Gamma/\gamma} =$  $+$ 

gets rid of the subdivergence originating in γ and $\int(\mathscr{C}[\omega_\Gamma] \otimes \omega_{\mathbb{I}}) = \int \mathscr{C}[\omega_\Gamma] =$ takes account of the overall divergence of the graph Γ.

The reader is encouraged to check for all the above example graphs that if one applies the coproduct again, it does not matter which side of the tensor product it acts on, ie

$$(\mathrm{id} \otimes \Delta) \circ \Delta(G) = (\Delta \otimes \mathrm{id}) \circ \Delta(G) \tag{4.1.36}$$

for any of the above Feynman graphs G and in fact, as proven in [**CoKrei00**], even for any element $G \in \mathcal{H}$. This property is called *coassociativity*. In terms of commutative diagrams, this identity takes the form

$$\begin{array}{ccc} \mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H} & \xleftarrow{\mathrm{id}\otimes\Delta} & \mathcal{H} \otimes \mathcal{H} \\ {\scriptstyle \Delta\otimes\mathrm{id}}\uparrow & & \uparrow{\scriptstyle \Delta} \\ \mathcal{H} \otimes \mathcal{H} & \xleftarrow{\Delta} & \mathcal{H} \end{array} . \tag{4.1.37}$$

The *associativity* of the product in $\mathcal{H}$ can also be described in terms of such a diagram, if we write it as a map $m\colon \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$, ie $m(\Gamma_1 \otimes \Gamma_2) := \Gamma_1\Gamma_2$. The commutative diagram then is given by

$$\begin{array}{ccc} \mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H} & \xrightarrow{\mathrm{id}\otimes m} & \mathcal{H} \otimes \mathcal{H} \\ \downarrow{\scriptstyle m\otimes\mathrm{id}} & & {\scriptstyle m}\downarrow \\ \mathcal{H} \otimes \mathcal{H} & \xrightarrow{m} & \mathcal{H} \end{array} , \tag{4.1.38}$$

which says nothing but

$$m(a \otimes m(b \otimes c)) = m(m(a \otimes b) \otimes c) \qquad \forall a, b, c \in \mathcal{H}. \tag{4.1.39}$$

4.1.3. Unit and counit. To arrive at a Hopf algebra, we need to introduce some more maps. One of them is the *unit map* $u\colon \mathbb{Q} \to \mathcal{H}$, which is given by $u(\lambda) = \lambda\mathbb{I}$, ie it maps onto the trivial subspace

$$\mathcal{H}_0 := \mathbb{Q}\mathbb{I}. \tag{4.1.40}$$

There is in fact a *grading*[2] on the algebra $\mathcal{H}$, given through the loop number of a Feynman graph,

$$\mathcal{H} = \bigoplus_{n\geq 0} \mathcal{H}_n \tag{4.1.41}$$

where $\mathcal{H}_n$ is the subspace of elements in $\mathcal{H}$ with grading degree $n \in \mathbb{N}$. This number is defined through a *grading operator* whose eigenspaces constitute the grading. We define the grading operator $Y\colon \mathcal{H} \to \mathcal{H}$ as follows. For a 1PI graph G, we have

$$Y(G) := |G|G, \tag{4.1.42}$$

ie the loop number is the eigenvalue of Y with respect to the eigenvector G. On a product of two elements $a, b \in \mathcal{H}$, it acts as $Y(ab) = Y(a)b + aY(b)$ and it is implemented as a linear map (therefore 'operator'). A map of this behaviour is referred to as a *derivation*[3]. These properties define Y on $\mathcal{H}$ unambigiously and one defines the subspace $\mathcal{H}_n$ by $a \in \mathcal{H}_n :\Leftrightarrow Y(a) = na$ and says that a is homogeneous of degree n. In the mathematics literature, the grading of an algebra usually needs to satisfy

$$m(\mathcal{H}_n \otimes \mathcal{H}_m) \subset \mathcal{H}_{n+m}, \tag{4.1.43}$$

which is clearly given in our case.

The *counit* $\varepsilon\colon \mathcal{H} \to \mathbb{Q}$ is a linear map such that $\varepsilon(h) = 1$ for $h = \mathbb{I}$ and vanishing otherwise. It has the property

$$(\varepsilon \otimes \mathrm{id}) \circ \Delta = \mathrm{id} = (\mathrm{id} \otimes \varepsilon) \circ \Delta, \tag{4.1.44}$$

which can be easily checked:

$$(\varepsilon \otimes \mathrm{id}) \circ \Delta(G) = \varepsilon(\mathbb{I}) \otimes G + \varepsilon(G) \otimes \mathbb{I} + \sum_{\gamma \in Q(G)} \varepsilon(\gamma) \otimes G/\gamma = \varepsilon(\mathbb{I}) \otimes G = 1 \otimes G = G, \tag{4.1.45}$$

and likewise for $(\mathrm{id} \otimes \varepsilon) \circ \Delta$. The counit is multiplicative, ie $\varepsilon(ab) = \varepsilon(a)\varepsilon(b)$ for all $a, b \in \mathcal{H}$ and

$$\Delta \circ u(\lambda) = \Delta(\lambda\mathbb{I}) = \lambda\mathbb{I} \otimes \mathbb{I} = (u \otimes u)(\lambda \otimes 1) \tag{4.1.46}$$

for all $\lambda \in \mathbb{Q}$. Due to $\mathbb{Q} \simeq \mathbb{Q} \otimes \mathbb{Q}$, this property is written as $\Delta \circ u = u \otimes u$ and expresses the compatibility of the unit u with the coproduct Δ.

With these properties of unit and counit, which are fulfilled for the above-described algebra of Feynman graphs and can be easily checked by the reader, the quintuple $(\mathcal{H}, m, \Delta, u, \varepsilon)$ is called a *bialgebra*, or more precisely, an associative and coassociative bialgebra with unit and counit[4].

The grading of a bialgebra must cohere with the coproduct, ie

$$\Delta(\mathcal{H}_n) \subset \bigoplus_{j=0}^{n} \mathcal{H}_j \otimes \mathcal{H}_{n-j}, \tag{4.1.47}$$

which is the case for Feynman graphs.

[2] See Appendix Section A.7 for an introduction to gradings.

[3] See Appendix A.7 for a definition.

[4] For more on these structures, see Appendix Section A.2.

4.1.4. Antipode. With these structures, we are able to introduce a group structure on $\mathcal{L}(\mathcal{H})$, the vector space of linear maps $\mathcal{H} \to \mathcal{H}$. If we take two such maps $f, g \in \mathcal{L}(\mathcal{H})$, then

$$(f \star g) := m \circ (f \otimes g) \circ \Delta \tag{4.1.48}$$

is again a map in $\mathcal{L}(\mathcal{H})$. This constitutes an associative bilinear operation on the set $\mathcal{L}(\mathcal{H})$ called *convolution*. In particular, this means for a 1PI graph Γ

$$(f \star g)(\Gamma) = \sum_{(\Gamma)} f(\Gamma')g(\Gamma'') = f(\mathbb{I})g(\Gamma) + f(\Gamma)g(\mathbb{I}) + \sum_{\gamma \in Q(\Gamma)} f(\gamma)g(\Gamma/\gamma). \tag{4.1.49}$$

Linearity of $h := f \star g$ is a consequence of the linearity of Δ, $f \otimes g$ and the product map m. On a nontrivial product of two 1PI graphs $G_1, G_2 \in \mathcal{H}$ this map evaluates to

$$(f \star g)(G_1 G_2) = \sum_{(G_1)} \sum_{(G_2)} f(G_1' G_2') g(G_1'' G_2'') \tag{4.1.50}$$

on account of (4.1.21). The bilinear operation $\star$ on $\mathcal{L}(\mathcal{H})$ is associative, ie $(f \star h) \star g = f \star (h \star g)$ and has in fact also a neutral element: the map $e := u \circ \varepsilon \colon \mathcal{H} \to \mathcal{H}$ is not just a projector onto $\mathcal{H}_0$ but also the neutral element of the convolution. Just set $g = e$ in (4.1.49), then

$$(f \star e)(\Gamma) = f(\Gamma) \tag{4.1.51}$$

and likewise $(e \star f)(\Gamma) = f(\Gamma)$ is obvious. Associativity is proven in Appendix Section A.3. One can now naturally define $\star$-powers by setting $f^{\star 0} := e$, $f^{\star 1} := f$ and $f^{\star n+1} := f \star f^{\star n}$, recursively. Even exponentials

$$\exp_\star(f) := \sum_{n \geq 0} \frac{f^{\star n}}{n!} \tag{4.1.52}$$

may exist. However, let us first see whether one can find a $\star$-inverse for a linear map f. For this to exist, we must make sure that the von Neumann series

$$f^{\star -1} = (e - (e - f))^{\star -1} = \sum_{n \geq 0} (e - f)^{\star n} \tag{4.1.53}$$

can be made sense of. This is not the case for all maps in $\mathcal{L}(\mathcal{H})$ but for those with $f(\mathbb{I}) = \mathbb{I}$, where the grading property (4.1.47) guarantees that the von Neumann series terminates on account of $(e - f)(\mathbb{I}) = 0$ (Appendix Section A.3 has a proof).

The *antipode* $S \in \mathcal{L}(\mathcal{H})$, sometimes called *coinverse*, is now defined by the identity

$$S \star \mathrm{id} = \mathrm{id} \star S = e. \tag{4.1.54}$$

First, we note that $\mathbb{I} = e(\mathbb{I}) = (S \star \mathrm{id})(\mathbb{I}) = S(\mathbb{I})\mathbb{I} = S(\mathbb{I})$ and because $\mathrm{id}(\mathbb{I}) = \mathbb{I}$ trivially, we know that the inverse of id, and hence the antipode exists, that is,

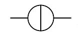
$$S := \sum_{n \geq 0} (e - \mathrm{id})^{\star n}. \tag{4.1.55}$$

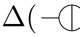
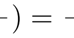
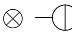
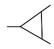
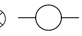
However, the grading makes sure that S is uniquely determined recursively. The recursive definition is as follows. We take a 1PI graph Γ and write

$$0 = e(\Gamma) = (S \star \mathrm{id})(\Gamma) = \sum_{(\Gamma)} S(\Gamma')\Gamma'' = S(\Gamma) + \Gamma + \sum_{\gamma \in Q(\Gamma)} S(\gamma)\ \Gamma/\gamma, \tag{4.1.56}$$

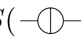
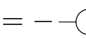
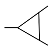
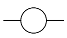
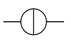
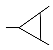
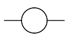
which implies the recursion

$$S(\Gamma) = -\Gamma - \sum_{\gamma \in Q(\Gamma)} S(\gamma)\ \Gamma/\gamma \qquad \text{(antipode)} \tag{4.1.57}$$

For a primitive graph this says, in particular $S(\Gamma) = -\Gamma$. Let us consider a non-primitive example, the graph $\Gamma =$ , for which the coproduct yields

(4.1.58) $\Delta($  $) =$  $\otimes \mathbb{I} + \mathbb{I} \otimes$  $+ 2$  $\otimes$  .

The recursion for the antipode then is

(4.1.59) $S($  $) = -$  $- 2\, S($  $)$  $= -$  $+ 2$   .

If $|\Gamma| = n$, then clearly $\Gamma/\gamma \in \bigoplus_{j=1}^{n-1} \mathcal{H}_j$ for all $\gamma \in Q(\Gamma)$. The antipode has the following properties:

(i) $S(ab) = S(b)S(a)$ for all $a, b \in \mathcal{H}$, ie S is an *algebra antimorphism.* Since the bialgebra of Feynman graphs $\mathcal{H}$ is a commutative algebra, this property means that it is multiplicative and hence also an algebra morphism;

(ii) $\tau(S \otimes S) \circ \Delta = \Delta \circ S$, where $\tau(a \otimes b) := b \otimes a$ is the flip map. This feature means that S is a *coalgebra antimorphism.*

(iii) $e \circ S = e$, which is easy to see: both sides acts as a projector onto $\mathcal{H}_0 = \mathbb{I}\mathbb{Q}$.

The proofs can be found in any book on Hopf algebras, a classical source is [**Sw69**]. For completeness and to summarise,

Definition 4.3 (Hopf algebra). *A Hopf algebra over $\mathbb{Q}$ is a hextuple $(\mathcal{H}, m, \Delta, u, \varepsilon, S)$ composed of an associative $\mathbb{Q}$-algebra $\mathcal{H}$ with*

(1) *product map $m\colon \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$ and unit $u : \mathbb{Q} \to \mathcal{H}$,*

(2) *multiplicative coproduct $\Delta\colon \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ and counit $\varepsilon\colon \mathcal{H} \to \mathbb{Q}$*

(3) *antipode $S\colon \mathcal{H} \to \mathcal{H}$, defined as the inverse of the identity map* id *on $\mathcal{H}$ with respect to the convolution product*

$$f \star g = m \circ (f \otimes g) \circ \Delta, \tag{4.1.60}$$

for linear maps $f, g : \mathcal{H} \to \mathcal{H}$, that is, $S \star \mathrm{id} = \mathrm{id} \star S = e$, where $e := u \circ \varepsilon$ is the neutral element of the convolution product (4.1.60).

$\mathcal{H}$ is called connected *if it has a grading $\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n$ such that $\mathcal{H}_0 = \mathbb{Q}\mathbb{I}$, where $u(1) = \mathbb{I}$ and both product and coproduct have the grading property*

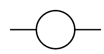
$$m(\mathcal{H}_n \otimes \mathcal{H}_m) \subset \mathcal{H}_{n+m}, \qquad \Delta(\mathcal{H}_n) \subset \bigoplus_{j=0}^{n} \mathcal{H}_j \otimes \mathcal{H}_{n-j}. \tag{4.1.61}$$

4.1.5. Hopf algebra of Feynman graphs. Finally, we have

Theorem 4.4. *The algebra $\mathcal{H}$ of 1PI Feynman graphs described above is a Hopf algebra, the Hopf algebra of Feynman graphs.*

Proof. The existence of the antipode is proven in Appendix Section A.3, Prop.A.6. The only thing left to prove is coassociativity of the coproduct. The reader is referred to [**CoKrei00**] for a proof of this property. □

There are some very simply examples of Hopf subalgebras of $\mathcal{H}$. The simplest is $\mathcal{H}_0 = \mathbb{Q}\mathbb{I}$. In fact, one can take a Feynman graph and use it as a generator of a Hopf algebra. A simple example can be constructed from the primitive graph

(4.1.62) $\gamma =$  .

The commutative unital $\mathbb{Q}$-algebra freely generated from this graph has a linear basis simply consisting of monomials γ^n, $n \geq 1$. The coproduct does not bring in anything new

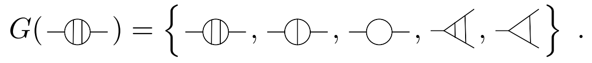
$$\Delta(\gamma) = \gamma \otimes \mathbb{I} + \mathbb{I} \otimes \gamma\ , \tag{4.1.63}$$

since the neutral element $\mathbb{I}$ is always tacitly assumed to be contained in the freely generated algebra. This already yields an infinite dimensional Hopf subalgebra $\mathcal{H}(\gamma) = \langle\gamma\rangle_{\mathbb{Q}}$ generated by just one primitive 1PI Feynman graph $\gamma \in \mathcal{G}$.

If a 1PI graph $\Gamma \in \mathcal{G}$ is not primitive and has subdivergences, we add these divergent subgraphs and their corresponding cographs to the generator set, denoted by $G(\Gamma)$, and obtain a Hopf algebra which we denote by $\mathcal{H}(\Gamma)$. Examples of generator sets are

(4.1.64) 

and

(4.1.65) 

These finitely generated Hopf algebras have a natural grading $\mathcal{H}(\Gamma) = \bigoplus_{n\geq 0} \mathcal{H}_n(\Gamma)$, given by the loop number, as in the case of the Hopf algebra of all 1PI Feynman graphs.

4.2. Feynman rules as Hopf algebra characters

Let now $\mathcal{H}$ be a connected Hopf algebra and $\mathcal{A}$ an associative and commutative algebra with neutral element $1_{\mathcal{A}} \in \mathcal{A}$. In the following, we consider algebra morphisms from $\mathcal{H}$ into the 'target algebra' $\mathcal{A}$. We assume that there exists a so-called *Rota-Baxter operator*, defined to be an operator $R\colon \mathcal{A} \to \mathcal{A}$ such that

$$R(ab) + R(a)R(b) = R(R(a)b + aR(b)) \tag{4.2.1}$$

for all $a, b \in \mathcal{A}$. To have something concrete and tangible in mind, one may imagine $\mathcal{A}$ to be an algebra of functions and R as an evaluation map that evaluates these functions at specific values of their arguments. Then, of course, (4.2.1) is trivially satisified.

4.2.1. Hopf algebra characters.

4.2.1. Hopf algebra characters. Let us next consider linear and multiplicative maps

$$\chi : \mathcal{H} \to \mathcal{A}\ , \qquad h \mapsto \chi(h) \tag{4.2.2}$$

that preserve the unit, ie $\chi(\mathbb{I}) = 1_{\mathcal{A}}$, so-called *Hopf algebra characters.* This property guarantees that they form a group with respect to the character convolution[5]

$$\chi \star \psi := m_{\mathcal{A}} \circ (\chi \otimes \psi) \circ \Delta, \tag{4.2.3}$$

in which $m_{\mathcal{A}}$ is the associative multiplication map of $\mathcal{A}$. The proof of this assertion can be found in Appendix Sections A.3,A.5. We denote the character group by $\mathrm{Ch}(\mathcal{H}, \mathcal{A})$ and the neutral element by e.

What we mean in particular are the assignments $\gamma \mapsto \int \omega_\gamma$ that we have discussed in the previous section. In their unregularised and unrenormalised form, of course, the integrals one obtains are only formal pairs which carry data about tempered distributions. But if we assume them to be regularised, then they evaluate to functions depending on the external momenta and the regulator $z \in \mathbb{C}$ (or cutoff). Because these functions generally have poles, for example at $z = 0$ (in dimensional regularisation), the target algebra is $\mathcal{A} = \mathbb{C}[z^{-1}, z]]$, ie the set of Laurent series with a finite number of pole terms.

Because the assignment of a Feynman graph γ to a Feynman integral $\int \omega_\gamma$ is part of the standard Feynman rules, we call the elements in $\mathrm{Ch}(\mathcal{H}, \mathcal{A})$ *Feynman characters.*

[5] We use the same sign as for the convolution on $\mathcal{L}(\mathcal{H})$. There should be no potential for confusion.

4.2.2. Renormalisation. Let $\chi \in \mathrm{Ch}(\mathcal{H}, \mathcal{A})$ be some Feynman character. We define a projector $P := \mathrm{id} - \mathrm{e}$ with

$$\mathrm{Aug} := \bigoplus_{n \geq 1} \mathcal{H}_n \tag{4.2.4}$$

as image, called *augmentation ideal.* It is indeed a *Hopf ideal* since the requirements

$$m(\mathcal{H} \otimes \mathrm{Aug}) = m(\mathrm{Aug} \otimes \mathcal{H}) \subset \mathrm{Aug}\ , \qquad \Delta(\mathrm{Aug}) \subset \mathrm{Aug} \otimes \mathcal{H} + \mathcal{H} \otimes \mathrm{Aug} \tag{4.2.5}$$

and $S(\mathrm{Aug}) \subset \mathrm{Aug}$ are satisfied[6]. We associate to the Feynman character χ the so-called *counterterm (character)* $S_R^\chi \in \mathrm{Ch}(\mathcal{H}, \mathcal{A})$, by requiring

$$S_R^\chi(G) = -R(S_R^\chi \star \chi P)(G) \tag{4.2.6}$$

for a 1PI divergent Feynman graph G, where $\chi P := \chi \circ P$ is a shorthand. It is not at all obvious that this map is a character, given that $\chi P(\mathbb{I}) = 0$, ie considering that $\chi P \notin \mathrm{Ch}(\mathcal{H}, \mathcal{A})$.

However, the Rota-Baxter condition ensures that this is still the case as long as the two characters χ and $S_R^\chi \star \chi$ furnish what is known as an *algebraic Birkhoff decomposition.* Whether or not this latter situation is given depends on the target algebra. For the cases we encounter in QFT, it is true (for details and a proof, see Appendix Section A.4).

The character $\chi_R \in \mathrm{Ch}(\mathcal{H}, \mathcal{A})$ defined by

$$\chi_R := S_R^\chi \star \chi \tag{4.2.7}$$

is called *renormalised Feynman character* associated with the character χ. Notice that $\psi\mathrm{e} := \psi \circ \mathrm{e}$ is the unique neutral element for all $\psi \in \mathrm{Ch}(\mathcal{H}, \mathcal{A})$ with respect to the $\star$-convolution product. We rewrite

$$\chi_R = S_R^\chi \star \chi = S_R^\chi \star [\chi(\mathrm{e} + P)] = S_R^\chi + S_R^\chi \star \chi P =: S_R^\chi + \overline{\chi}, \tag{4.2.8}$$

where $\overline{\chi}$ is called *Bogoliubov character.* This character describes the assigment of a Feynman graph to a Feynman integral which has been cured of its subdivergences: if we use (4.2.6), then

$$\chi_R = S_R^\chi + \overline{\chi} = -R(S_R^\chi \star \chi P) + \overline{\chi} = (\mathrm{id} - R)\overline{\chi} \tag{4.2.9}$$

represents the last subtraction in which the overall divergence is removed.

4.2.3. Feynman rules as characters. We shall now be more concrete about these characters and, in particular, the target algebra $\mathcal{A}$. In Section 4.1, we have written Feynman integrals as differential forms $\int \omega_G$ associated with a 1PI graph G. As alluded to, the assignment of Feynman graphs to these (divergent) integrals is represented by a Feynman character,

$$G \mapsto \chi(G) := \int \omega_G \tag{4.2.10}$$

in which we have left the regularisation implicit. As we have discussed in the previous section, once renormalised, these integrals are generalised functions in momentum space and may be applied to Schwartz functions to yield a complex number $\chi_R(G, f) = \int \mathscr{R}[\omega_G](f)$.

4.2.4. Renormalisation of Feynman characters. Let us see how the above formalism works. Assume G is a primitive divergent 1PI graph, ie a graph with no subdivergence: $\Delta(G) = \mathbb{I} \otimes G + G \otimes \mathbb{I}$. Then

$$\begin{aligned}
\chi_R(G) &= (S_R^\chi \star \chi)(G) = (S_R^\chi \otimes \chi) \circ \Delta(G) = S_R^\chi(\mathbb{I})\chi(G) + S_R^\chi(G)\chi(\mathbb{I}) = \chi(G) + S_R^\chi(G) \\
&= \chi(G) - R(S_R^\chi \star \chi P)(G) = \chi(G) - R(\chi(G)) = \int \omega_G - R \int \omega_G \\
&= \int (\mathrm{id} - R)\omega_G = \int (\omega_G + \mathscr{C}(\omega_G)) = \int \mathscr{R}[\omega_G],
\end{aligned} \tag{4.2.11}$$

[6]For a concise introduction to ideals, see Appendix Section A.6.

where, for scalar theories, the Rota-Baxter operator $R\omega_G = -\mathscr{C}(\omega_G)$ is defined as a Taylor expansion map which yields a Taylor expansion of the integrand up to order $D(G) \geq 0$ with respect to the external momenta at a fixed reference point.

It does not matter whether this is done before or after the integration, as long as the regularisation is in place. The case of gauge theories is more subtle as the Feynman integrals there, for example in QED, are matrix-valued. The details for QED are described in [**Sui06**].

Let now Γ be 1PI with only one 1PI divergent subgraph γ. The renormalised Feynman character evaluates this to

$$\chi_R(\Gamma) = S_R^\chi(\Gamma) + \chi(\Gamma) + S_R^\chi(\gamma)\chi(\Gamma/\gamma) = S_R^\chi(\Gamma) + \overline{\chi}(\Gamma). \tag{4.2.12}$$

First we compute $S_R^\chi(\gamma) = -R(S_R^\chi \star \chi P)(\gamma) = -R(\chi(\gamma)) = -\int R\omega_\gamma$. The Bogoliubov character is therefore given by

$$\overline{\chi}(\Gamma) = (S_R^\chi \star \chi P)(\Gamma) = \chi(\Gamma) + S_R^\chi(\gamma)\chi(\Gamma/\gamma) = \int (\omega_\Gamma - R(\textstyle\int \omega_\gamma)\omega_{\Gamma/\gamma}). \tag{4.2.13}$$

We see that it cures the Feynman integral $\chi(\Gamma) = \int \omega_\Gamma$ of its subdivergence $\chi(\gamma) = \int \omega_\gamma$. Making use of (4.2.13), we see that the counterterm character S_R^χ provides the overall subtraction when acted on Γ:

$$\begin{aligned} S_R^\chi(\Gamma) &= -R(S_R^\chi \star \chi P)(\Gamma) = -R\int \{\omega_\Gamma - R(\textstyle\int \omega_\gamma)\omega_{\Gamma/\gamma}\} \\ &= \int \{-R(\omega_\Gamma) + R(\omega_\gamma) \otimes R(\omega_{\Gamma/\gamma})\} =: \int \mathscr{C}(\omega_\Gamma), \end{aligned} \tag{4.2.14}$$

which then leads to

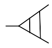
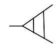
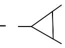
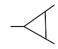
$$\begin{aligned} \chi_R(\Gamma) = (\mathrm{id} - R)\overline{\chi}(\Gamma) &= \int (\mathrm{id} - R)(\omega_\Gamma - R(\textstyle\int \omega_\gamma)\omega_{\Gamma/\gamma}) \\ &= \int \{\omega_\Gamma - R(\textstyle\int \omega_\gamma)\omega_{\Gamma/\gamma} - R(\omega_\Gamma) + R(\textstyle\int \omega_\gamma) R(\omega_{\Gamma/\gamma})\} \\ &= \int \{\omega_\Gamma \underbrace{-R(\omega_\gamma)}_{=\mathscr{C}(\omega_\gamma)} \otimes \omega_{\Gamma/\gamma} \underbrace{-R(\omega_\Gamma) + R(\omega_\gamma) \otimes R(\omega_{\Gamma/\gamma})}_{=\mathscr{C}(\omega_\Gamma)}\} =: \int \mathscr{R}[\omega_\Gamma]. \end{aligned} \tag{4.2.15}$$

where the renormalised integrand is $\mathscr{R}[\omega_\Gamma] = \omega_{\mathbb{I}} \otimes \omega_\Gamma + \mathscr{C}(\omega_\gamma) \otimes \omega_{\Gamma/\gamma} + \mathscr{C}(\omega_\Gamma) \otimes \omega_{\mathbb{I}}$ in which the coproduct

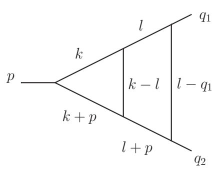
$$\Delta(\Gamma) = \mathbb{I} \otimes \Gamma + \Gamma \otimes \mathbb{I} + \gamma \otimes \Gamma/\gamma \tag{4.2.16}$$

is clearly visible. The overall counterterm $\mathscr{C}(\omega_\Gamma) = -R(\omega_\Gamma) + R(\omega_\gamma) \otimes R(\omega_{\Gamma/\gamma})$ reflects the fact that Γ still has an overall divergence, even after subtraction of its subdivergence.

Example: massless scalar vertex graph. To see a concrete example, we take the graph $\Gamma =$  in massless $(\varphi^3)_6$ theory. The coproduct yields

(4.2.17) $\Delta(\Gamma) =$  $\otimes\ \mathbb{I} + \mathbb{I}\ \otimes$  $+$  $\otimes$ 

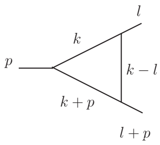
and the corresponding Feynman integral reads

(4.2.18) 

$$= \int \frac{d^d l}{l^2(l-q_1)^2(l+p)^2} \int \frac{d^d k}{k^2(k-l)^2(k+p)^2} = \int \omega_\Gamma(q_1, p)$$

Assume we regularize it by setting $d = 6 - 2z$. Then we first take care of the subintegration, ie the Feynman graph subsector

(4.2.19)

$\mathscr{R}($  $)$

$$= \int \frac{d^dk}{k^2(k-l)^2(k+p)^2} - \int \frac{d^dk}{k^2(k-l)^2(k+p)^2}\bigg|_{l^2=p^2=\mu^2}$$

$$= \int [\omega_\gamma(l,p) - \omega_\gamma(l_0,p_0)] = \int (\mathrm{id} - R)[\omega_\gamma](l,p) = \int \mathscr{R}[\omega_\gamma](l,p)$$

where $l, p \in \mathbb{R}^6$ are the external parameters and $l_0, p_0 \in \mathbb{R}^6$ the reference momentua, here chosen such that $l_0^2 = p_0^2 = \mu^2$. This implies the renormalisation condition $\mathscr{R}[\omega_\gamma](l_0,p_0) = 0$. We replace the subintegral in (4.2.18) by this term and get

(4.2.20)

$$I_d = \int \frac{d^dl}{l^2(l-q_1)^2(l+p)^2}\left(\int \frac{d^dk}{k^2(k-l)^2(k+p)^2} - \int \frac{d^dk}{k^2(k-l)^2(k+p)^2}\bigg|_{l^2=p^2=\mu^2}\right) .$$

$$= \int \{\omega_\Gamma - R[\textstyle\int\omega_\gamma]\omega_{\Gamma/\gamma}\}(q_1,p)$$

However, this is not an expression for which the limit $d \to 6$ exists on account of the logarithmic divergence of the l-integration. We need yet another subtraction to achieve this aim, that is, we must add the term

(4.2.21)

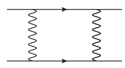
$$-RI_d =$$

$$-\left[\int \frac{d^dl}{l^2(l-q_1)^2(l+p)^2}\left(\int \frac{d^dk}{k^2(k-l)^2(k+p)^2} - \int \frac{d^dk}{k^2(k-l)^2(k+p)^2}\bigg|_{l^2=p^2=\mu^2}\right)\right]\Bigg|_{p^2=q_1^2=\mu^2}$$

$$= \int \{-R(\omega_\Gamma) + R(\textstyle\int\omega_\gamma)R(\omega_{\Gamma/\gamma})\}$$

and the physical limit $I_R = \lim_{d\to 6}\mathscr{R}(I_d) = \lim_{d\to 6}(I_d - RI_d)$ exists and, if we piece (4.2.20) and (4.2.21) together, we see that it is an instance of (4.2.15).

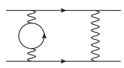
4.2.5. Overall convergent graphs. In the case that a graph G is overall convergent, ie $D(G) < 0$, we set $S_R^\chi(G) = 0$. If this graph has no subdivergences, like

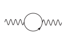
(4.2.22) $G =$  ,

then there is no need for renormalisation. One therefore demands $\chi_R(G) = \chi(G)$. This works out fine:

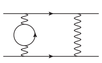
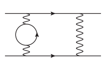
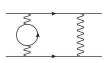
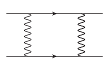
(4.2.23)
$$\chi_R(G) = (S_R^\chi \star \chi)(G) = S_R^\chi(\mathbb{I})\chi(G) + S_R^\chi(G)\chi(\mathbb{I}) = \chi(G) = \int \omega_G.$$

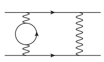
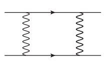
In case it has a subdivergence, like the graph

(4.2.24) $\Gamma =$  ,

with subdivergence $\gamma =$  , we find

(4.2.25)

$\chi_R($  $) = S_R^\chi(\mathbb{I})\chi($  $) + S_R^\chi($  $)_{=0}\,\chi(\mathbb{I}) + S_R^\chi($  $)\chi($  $)$

$= \chi($  $) - R[\chi($  $)]\chi($  $) = \int [\omega_\Gamma - R(\int\omega_\gamma)\omega_{\Gamma/\gamma}].$

As we have seen in Section 4.1, the renormalised integrand can be written as

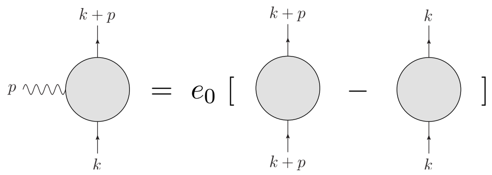
$$\mathscr{R}[\omega_\Gamma] = \mathscr{C}(\omega_{\mathbb{I}}) \otimes \omega_\Gamma + \mathscr{C}(\omega_\Gamma) \otimes \omega_{\mathbb{I}} + \mathscr{C}(\omega_\gamma) \otimes \omega_{\Gamma/\gamma}, \tag{4.2.26}$$

where $\mathscr{C}(\omega_\Gamma) = 0$ because $D(\Gamma) = -2 < 0$, ie Γ is overall convergent.

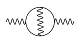
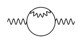
4.3. Hopf ideal and Ward identity in QED

The well-known gauge symmetry of quantum electrodynamics (QED) entails the so-called *Ward* or *Ward-Takahashi identity*, diagrammatically,

(4.3.1)



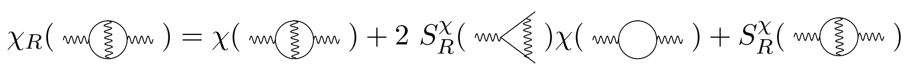
in which all blobs are full Green's functions, ie the vertex amplitude is non-amputated while the fermion propagators are full ones [**PeSch95**]. Courtesy of this identity, the wave function renormalisation constants of the fermion spinors and the charge can be choosen to be equal, ie $Z_1 = Z_2$ (charge renormalisation = electron wave renormalisation) [**Wa50**]. We shall now see how this manifests itself combinatorially in the form of a *Hopf ideal* in the Hopf algebra of Feynman diagrams $\mathcal{H}$ in QED [**Sui06**].

Two-loop example. We consider the three graphs $\Gamma_1 =$  $, \Gamma_2 =$  $, \Gamma_3 =$  , which will do us good service to motivate this Hopf ideal. In the Hopf-algebraic approach, one applies the renormalized Feynman character χ_R to the sum of these 3 graphs,

$$\chi_R(\Gamma_1+\Gamma_2+\Gamma_3) = (S_R^\chi \star \chi)(\Gamma_1+\Gamma_2+\Gamma_3) = (S_R^\chi \star \chi)(\Gamma_1) + (S_R^\chi \star \chi)(\Gamma_2) + (S_R^\chi \star \chi)(\Gamma_3) \tag{4.3.2}$$

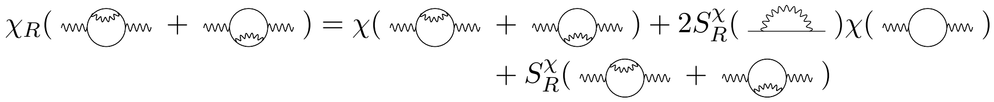
For the first graph we get

(4.3.3)



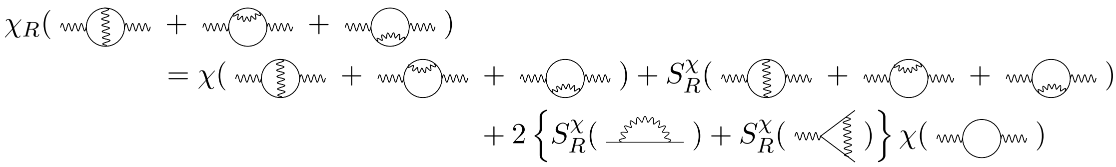
and

(4.3.4)



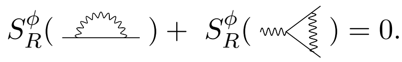
for the other two[7] Altogether, one finds

(4.3.5)



Thanks to the Ward identity (4.3.1), we can choose a renormalization scheme in which the term in curly brackets vanishes:

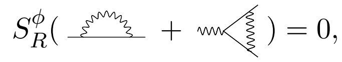
(4.3.6)



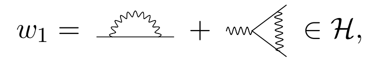
This simplyfies the renormalization procedure significantly: the sum of the three graphs need only one subtraction for the overall divergence, it behaves like a primitive element (ie a divergent graph void of subdivergences). Although the individual counterterms are needed for curing the subdivergences, their service becomes obsolete when we take the sum of the three graphs.

[7]We hope that readers formerly unfamiliar with the Hopf algebra of Feynman diagrams can by now confirm these results as an exercise.

4.3.1. Ward elements. If we choose the counterterm character S_R^χ such that (4.3.6) holds, we can on account of linearity write

(4.3.7) 

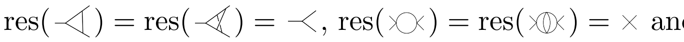
which means that the 1-loop *'Ward element'*, defined by

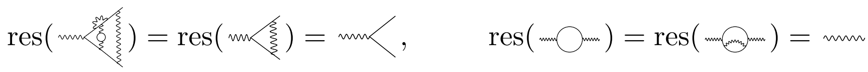
(4.3.8) 

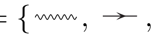
lies in the kernel of the counterterm character S_R^χ. Because of the different external leg structures of both graphs, an element like this one will never appear in the argument of χ or χ_R. For this reason, we may get the idea of equating all such Ward elements to zero in the first place.

Before we carry on elaborating on this idea, let us introduce some notation which shall be also useful in the next chapter on Dyson-Schwinger equations.

Definition 4.5 (Residue). *The* residue *of a graph* Γ *is the graph* $\mathrm{res}(\Gamma)$ *obtained from* Γ *by shrinking all internal edges to a single point.*

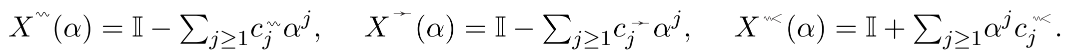
Examples are  and

(4.3.9) 

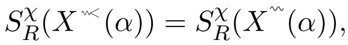
By $\mathcal{R}$ we denote the set of such 'residues of interest' for a given renormalisable theory. For QED, this set consists of only 3 elements, . What we mean by 'of interest' is that once one knows the 1PI perturbation series with these residues, ie the corresponding amplitudes, all other Green's functions can be pieced together from these 1PI series. We write the combinatorial 1PI series of a renormalisable QFT with one coupling constant in the form

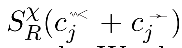
(4.3.10) $$X^r(\alpha) = \mathbb{I} \pm \sum_{j\geq 1} c_j^r \alpha^j\ , \qquad r \in \mathcal{R},$$

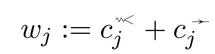
where $+$ is chosen for a vertex and $-$ for a propagator series. In QED, we shall denote these series by

(4.3.11) 

These series are elements of $\mathcal{H}[[\alpha]$, the set of formal power series in the coupling parameter α with coefficients in the Hopf algebra of QED Feynman graphs $\mathcal{H}$. The counterterm Ward identity $Z_1 = Z_2$ takes the form

(4.3.12) 

which means  for each loop order $j \geq 1$. The ideal is now constructed as follows. Because the Ward elements

(4.3.13) 

are mapped to zero by the counterterm character S_R^χ, so is anything in $\mathcal{H}$ of the form

(4.3.14) $$a = \sum_{j=1}^n b_j w_j \in \mathcal{H}$$

with some $b_j \in \mathcal{H}$, $(j = 1, ..., n)$, that is, $S_R^\chi(a) = 0$. The set

(4.3.15) $$\mathcal{I} = \sum_{j\geq 1} \mathcal{H} w_j$$

is an ideal since $\mathcal{H}\mathcal{I} = \mathcal{I}\mathcal{H} \subset \mathcal{I}$. This gives rise to an equivalence relation on the Hopf algebra of QED Feynman graphs: two elements $h, h' \in \mathcal{H}$ are equivalent if their difference lies in $\mathcal{I}$, ie we have

(4.3.16) $$h \sim h' \qquad \Leftrightarrow \qquad h = h' + a \qquad a \in \mathcal{I}\ .$$

To see that this ideal is Hopf, we have to draw on a result by Suijlekom in [**Sui07**]:

Proposition 4.6 (Coproduct formula). *For the combinatorial 1PI series X^r of residue r in QED, the coproduct yields*

$$\Delta(X^r) = X^r \otimes \mathbb{I} \pm \sum_{n \geq 1} X^r Q^n \otimes c_n^r \alpha^n, \tag{4.3.17}$$

where '+' is for a vertex and '−' for a propagator series and the so-called 'invariant charge' $Q \in \mathcal{H}[[\alpha]]$ is the formal series given by

$$Q := \frac{(X^{\sim\!<})^2}{X^{\sim}(X^{\rightarrow})^2} \ . \tag{4.3.18}$$

With this result at hand, the assertion that the ideal $\mathcal{I}$ is a co-ideal, ie $\Delta(\mathcal{I}) \subset \mathcal{I} \otimes \mathcal{H} + \mathcal{H} \otimes \mathcal{I}$ and that the antipode respects it, $S(\mathcal{I}) \subset \mathcal{I}$ is a straightforward corollary [**Sui07**]:

Corollary 4.7. *The ideal $\mathcal{I}$ is Hopf in $\mathcal{H}$, that is,*

$$\Delta(\mathcal{I}) \subset \mathcal{I} \otimes \mathcal{H} + \mathcal{H} \otimes \mathcal{I} \ , \qquad S(\mathcal{I}) \subset \mathcal{I} \tag{4.3.19}$$

and $\mathcal{I} \subset \ker \varepsilon$.

Proof. On account of both maps linearity and multiplicativity, it suffices to show $\Delta(w_j) \in \mathcal{I} \otimes \mathcal{H} + \mathcal{H} \otimes \mathcal{I}$ and $S(w_j) \in \mathcal{I}$ for an arbitrary Ward element. The series of Ward elements is given by

$$W = X^{\sim\!<} - X^{\rightarrow} = \sum_{n \geq 1} (\ c_n^{\sim\!<} + c_n^{\rightarrow} \ ) \alpha^n. \tag{4.3.20}$$

If we apply (4.3.17) to it, we get

$$\Delta(W) = \sum_{n \geq 1} \underbrace{[\ W Q^n \otimes c_n^{\sim\!<} + X^{\rightarrow} Q^n \otimes w_n \ ]}_{\in (\mathcal{I} \otimes \mathcal{H} + \mathcal{H} \otimes \mathcal{I})[[\alpha]]} \alpha^n. \tag{4.3.21}$$

For the antipode we first note that $S(w_1) = -w_1$ and then by induction

$$S(w_n) = -w_n - \sum_{(w_n)'} S(w_n') w_n'' \tag{4.3.22}$$

in which $\sum_{(x)'} x' \otimes x'' := \Delta(x) - x \otimes \mathbb{I} - \mathbb{I} \otimes x$ for $x \in \mathcal{H}$ is the 'reduced coproduct'. The last assertion is trivial by definition of the counit: it vanishes on any nontrivial element, only on the trivial subspace $\mathcal{H}_0 = \mathbb{Q}\mathbb{I}$ does it not vanish. □

What virtue this brings for renormalisation can be seen as follows: by taking the quotient

$$\mathcal{H}_\sim := \mathcal{H}/\mathcal{I}, \tag{4.3.23}$$

we obtain $W = 0$ and hence

$$\frac{X^{\sim\!<}}{X^{\rightarrow}} = \mathbb{I} \tag{4.3.24}$$

which entails $Q = \mathbb{I}/X^{\sim}$ for the invariant charge.

CHAPTER 5

Dyson-Schwinger equations and the renormalisation group

Dyson-Schwinger equations (DSEs) are integral equations that describe the relations between the different Green's functions of a QFT. Going back to the work of Dyson and Schwinger [**Dys49b, Schwi51**], there are two ways to obtain them.

The most intuitive one, in Dyson's spirit, as described in [**BjoDre65**], is suggested by the self-similiarity of Feynman diagram series. Although this approach is perturbative, it leads to the DSEs as equations that can rightfully be interpreted as nonperturbative equations.

The other path to DSEs, in Schwinger's spirit, makes use of functional integrals and is hence more technical. On the upside, however, it reveals the DSEs' origin as Euler-Lagrange equations of the theory.

In this work, we shall adopt the viewpoint of perturbation theory and propose that the DSEs, read as nonperturbative equations, be utilised to *define a quantum field theory*. As we shall see along the way, there are nontrivial examples which give us confidence that this route is viable.

Section 5.1 introduces DSEs based on the self-similiarity of Feynman diagram series by using the rainbow series as a pedagogical paradigm and expounds its purely combinatorial version as a fixed point equation in the algebra $\mathcal{H}[[\alpha]]$ of formal power series with Hopf algebra-valued coefficients.

The next section, Section 5.2, is devoted to Yukawa theory and reviews the standard approximations with increasing complexity and thereby explains a method which makes use of so-called *Mellin transforms*. We first discuss *linear DSEs*, namely the *rainbow* and *ladder approximations* [**Krei06**] whose DSEs both have the same form. Including a brief account of the *next-to-ladder approximation* [**BiKreiW07**], we describe a more general form of such DSEs and their solutions [**Ki12**].

It turns out that the anomalous dimensions are in the linear case generally algebraic functions of the coupling which entails that their Taylor series have a nonzero radius of convergence. We believe this feature to be unphysical.

In contrast to these linear examples, the so-called *Kilroy approximation*, an example of a (highly) *non-linear DSE*, turns out to have an anomalous dimension with a divergent perturbation series [**BroK01**]. Unfortunately, the corresponding DSE cannot be solved by the same Mellin transform method as it is the case for linear DSEs.

Section 5.3 introduces the DSEs of QED and their general form for theories of a single coupling parameter. Needless to say, they cannot be solved to this day. We finally discuss the DSEs of QED in the quotient Hopf algebra where some things must be modified in order for the DSEs to be still valid.

However, as Section 5.4 explains in the case of single-scale amplitudes, the combinatorial description of renormalisation entails the renormalisation group (RG) equation and a recursion formula for what we will call the amplitude's *log-coefficient functions* or simply *RG functions* [**KrY06**].

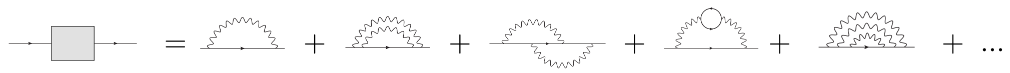
The necessary mathematical machinery of Hopf algebra characters and their Lie generators is introduced along the way. This material is by now standard and can be found in many places, for example [**Man04, EGraPa07**].

Following this, Section 5.5 reviews how to reformulate DSEs in terms of the Mellin transforms of the primitive Dyson-Schwinger skeletons first by using the simple examples from Yukawa theory and then finally for QED [**KrY06, Y11**]. In the case of the Kilroy approximation, this variant of the DSE implies a non-linear ordinary differential equation with a particularly simple 'inhomogeneity' which is derived in Section 5.6, before the chapter closes with the general version of this equation [**KrY06, BroK01**].

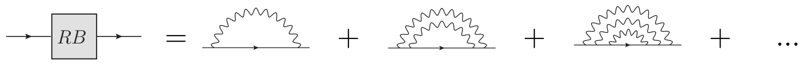
For the convenience of the reader, this chapter serves as a pedagogical introduction to Dyson-Schwinger equations in QFT and expounds the Hopf-algebraic combinatorial approach in a way that tries to be as comprehensible as possible. We have selected only those topics we find absolutely necessary to understand the two subsequent chapters. Topics such as Hochschild cohomology, combinatorial Dyson-Schwinger equations in the Hopf algebra of words or decorated rooted trees and the connection to number theory are omitted altogether. The interested reader is referred to the original literature [**Krei06a, Foi10**] or the lecture notes [**KlaKrei13a, KlaKrei13b**].

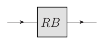
5.1. Combinatorial and analytic Dyson-Schwinger equations

5.1.1. Self-similiarity of Feynman diagram series. As a warm-up, let us consider a simple example to illustrate the combinatorial approach and see how it yields a non-perturbative toy model DSE. In the following, we denote by a shaded box the fermion self-energy of QED, ie the 1PI Green's function:

(5.1.1) 

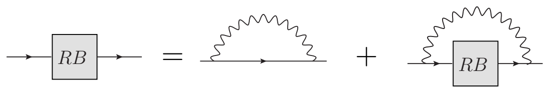
Consider the so-called *rainbow approximation* which is depicted by the series

(5.1.2) 

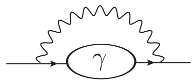
If we write this schematically in terms of integrals with $\int \omega =$  and $\Sigma_{RB} =$ , then

$$\begin{aligned}\Sigma_{RB} &= \int \omega \; + \int \omega \int \omega \; + \int \omega \int \omega \int \omega \; + \; ... \\ &= \int \omega \; (1 + \int \omega \; + \int \omega \int \omega \; + ...) = \int \omega \; (1 + \Sigma_{RB}),\end{aligned} \tag{5.1.3}$$

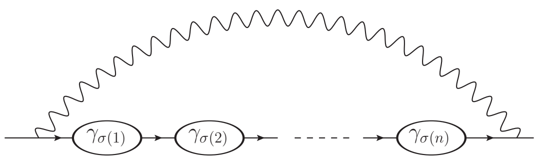
Diagrammatically, this is

(5.1.4) 

For a combinatorial description, as introduced in Section 4.3, we define a linear *insertion operator* $B_+^{⌓}$ on the Hopf subalgebra $\mathcal{H}_{RB}$ generated by all rainbow graphs in (5.1.2) by setting

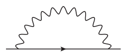
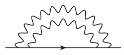
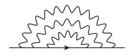
(5.1.5) $B_+^{⌓}(\gamma) :=$ 

for a rainbow graph γ and for a product of rainbow graphs $\gamma_1...\gamma_n$, we define

(5.1.6) $B_+^{⌓}(\gamma_1...\gamma_n) := \dfrac{1}{n!} \sum_{\sigma \in S_n}$  , where S_n is the set of permutations of the elements in $\{1, 2, ..., n\}$, ie the symmetric group. Note that the meaning of the graph  as upper index is that of a *skeleton graph* into which the insertion operator inserts whatever it is given as an argument. This operator enables us to write the rainbow DSE (5.1.4) for the rainbow series as a formal power series, ie an element in $\mathcal{H}_{RB}[[\alpha]]$:

$$X(\alpha) = \mathbb{I} + \alpha B_+^{\frown}(X(\alpha)), \tag{5.1.7}$$

where the solution to this DSE is the formal series $X(\alpha)$ is given by

(5.1.8) $X(\alpha) = \mathbb{I} +$  $\alpha +$  $\alpha^2 +$  $\alpha^3 + \quad ...$

Note that the rainbow DSE (5.1.7) never creates a situation in which the insertion operator $B_+^{\frown}$ is confronted with a nontrivial product of rainbow graphs. For this reason, the definition in (5.1.6) is not necessary for the formulation of the rainbow DSE: as (5.1.6) shows, the rainbow Hopf algebra $\mathcal{H}_{RB}$ is not closed under the action of the insertion operator.

5.1.2. Combinatorial versus analytic DSEs. We would like to point out again that the combinatorial stance differs slightly from the conventional one. When physicists draw diagrammatic expressions like the series (5.1.1) or (5.1.2) they really mean the corresponding series of Feynman integrals and tacitly never view Feynman diagram themselves as algebraic objects, ie the DSE (5.1.4) is usually seen as a shorthand for

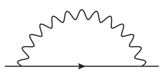
$$- i\Sigma_{RB}(q) = \alpha \int \frac{d^4k}{4\pi^3} \gamma^\mu S_0(k)[1 + \Sigma_{RB}(k)S_0(k)]\gamma^\nu D^0_{\mu\nu}(q - k), \tag{5.1.9}$$

with the obvious conventional notation (downstairs and upstairs '0' stand for 'free'). We call such equation *analytic* DSE and by adopting the combinatorial approach, we take (5.1.8) seriously as the algebraic solution of the *fixed point equation* (5.1.7) in the formal algebra $\mathcal{H}_{RB}[[\alpha]]$ and strictly distinguish between *combinatorial* and *analytic* DSEs.

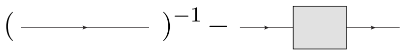
To pass from the combinatorial DSE (5.1.7) to its analytic version (5.1.9), we employ the very Hopf algebra characters that we have introduced in the preceding chapter, Section 4.2. With one subtlety though. Combinatorial DSEs for a proper renormalisable QFT, and we will discuss their general form and in particular those of QED in due course, require us to use the unit element $\mathbb{I} \in \mathcal{H}$. The reason lies in the relation between the Feynman graph series representing the *full* propagator and that representing the self-energy, ie the propagator's 1PI series. Take the rainbow series for instance. As a solution of (5.1.7) the series

$$\widetilde{X}(\alpha) = \frac{\mathbb{I}}{\mathbb{I} - [X(\alpha) - \mathbb{I}]} = \mathbb{I} + \sum_{m \geq 1} [X(\alpha) - \mathbb{I}]^m = \mathbb{I} + \sum_{m \geq 1} \sum_{k=0}^{m} \binom{m}{k} (-1)^k X(\alpha)^k \tag{5.1.10}$$

describes the full propagator series in the rainbow approximation. This identity is (mutatis mutandis) a generic one for 1PI propagator series and their associated connected series. No matter whether we include $\mathbb{I}$ in the self-energy series, the series corresponding to the full propagator must start with the neutral element $\mathbb{I}$. This is the element which gets mapped by Hopf algebra characters in $\mathrm{Ch}(\mathcal{H}, \mathcal{A})$ to the neutral element $1_{\mathcal{A}}$ of the target algebra (see Section 4.2). We therefore define Feynman characters in such a way that they map a Feynman graph to a form factor. For example, consider the form factor decomposition

(5.1.11)  $= \not{q} A_1(q^2, m^2) + m B_1(q^2, m^2)$

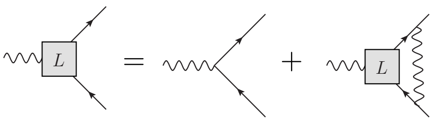
with two form factors $A_1(q^2, m^2)$ and $B_1(q^2, m^2)$, where the index refers to '1-loop'. For the inverse full propagator, one finds

(5.1.12)  $= \not{q}[1 - A_1(q^2, m^2)\alpha + \ldots] + m[1 - B_1(q^2, m^2)\alpha + \ldots]\ .$

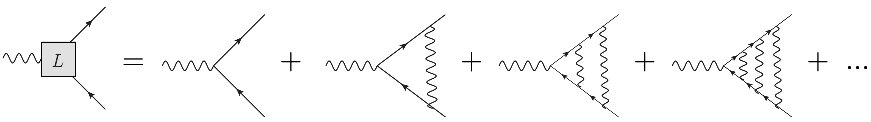
This means in particular that we have to split the self-energy series into two parts each of which requires an extra DSE. This leads even for the rainbow approximation to a system of two coupled DSEs in our formalism. Since the results we present in the last section of this chapter have been obtained for the massless case, we shall avoid the complications brought about by the electron's mass and consider only the massless case in this work.

5.2. Approximations in Yukawa theory

Both the rainbow DSE (5.1.4) and the 'ladder' DSE

(5.2.1) 

for the so-called *ladder approximation*

(5.2.2) 

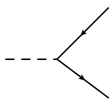
have so far only been solved in massless Yukawa and massless scalar $(\varphi^3)_6$ theory[1] (at zero momentum transfer for the ladder series). These results have been attained in the late 1990s using dimensional regularisation, where the solutions turned out to be of the same form [**DeKaTh96, DeKaTh97**].

We will briefly rederive their results by using another technique, the *method of Mellin transforms*, as introduced by Kreimer and Yeats in [**Krei06, KrY06**]. In preparation for this method, we first introduce the *Mellin transform* of a primitive 1PI graph.

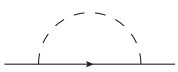
5.2.1. Mellin transform of a primitive graph. Let us remind ourselves of Yukawa theory and its Feynman rules. In the massless case, this theory is characterised by the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi + i\overline{\psi}\not{\partial}\psi - g\varphi\overline{\psi}\psi \tag{5.2.3}$$

and describes massless spin one-half fermions and scalar mesons represented by the spinor field ψ and the scalar field φ, respectively. The Feynman rules in momentum space are

(5.2.4)  $= -ig\ ,$ $\quad p$ $= \dfrac{i}{\not{p} + i\epsilon}\ ,$ $\quad p$ $= \dfrac{i}{p^2 + i\epsilon}$

accompanied with the corresponding integration directives. Before we come to the rainbow DSE in Yukawa theory, let us compute

(5.2.5) q  $= (-ig)^2 \displaystyle\int \frac{d^4k}{(2\pi)^4} \frac{i}{\not{k} + i\epsilon} \frac{i}{(q-k)^2 + i\epsilon} = \not{q} A_1(q^2)$

[1] What sets QED apart from the Yukawa case is the Lorentz tensor structure of the photon propagator, making it all the more harder to obtain the corresponding results.

Performing the standard steps to extract the form factor $A_1(q^2)$, and going Euclidean by setting $k_4 = -ik_0$ as well as $q_4 = -iq_0$, we obtain

$$A_1(-q_E^2) = \frac{ig^2}{2} \int \frac{d^4 k_E}{(2\pi)^4} \frac{1}{k_E^2 (q_E - k_E)^2} \tag{5.2.6}$$

in which the index 'E' stands for Euclidean, that is, $k_E = (k_1, k_2, k_3, k_4)$ (and likewise with q_E). We regularise it by means of a convergence factor ('analytical regularisation')

$$(k_E^2)^{-\rho} \tag{5.2.7}$$

with regulator $\rho \in \mathbb{C}$ and use the well-known 'master formula'

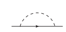
$$\int \frac{d^4 k_E}{(2\pi)^4} \frac{1}{(k_E^2)^r ((k_E - q_E)^2)^s} = (q_E^2)^{-(r+s-2)} \frac{1}{(4\pi)^2} \frac{\Gamma(r+s-2)\Gamma(2-r)\Gamma(2-s)}{\Gamma(r)\Gamma(s)\Gamma(4-r-s)}, \tag{5.2.8}$$

to get

$$\begin{aligned} A_1^\rho(-q_E^2) &= \frac{ig^2}{2} \int \frac{d^4 k_E}{(2\pi)^4} \frac{(k_E^2)^{-\rho}}{k_E^2 (k_E - q_E)^2} = \frac{ig^2}{2} (q_E^2)^{-\rho} \frac{1}{(4\pi)^2} \frac{\Gamma(\rho)\Gamma(1-\rho)}{\Gamma(1+\rho)\Gamma(2-\rho)} \\ &= \frac{ig^2}{2(4\pi)^2} (q_E^2)^{-\rho} \frac{1}{\rho(1-\rho)} =: ia\ (q_E^2)^{-\rho} F(\rho). \end{aligned} \tag{5.2.9}$$

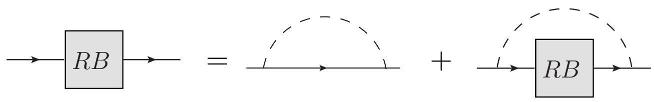
By defining the new coupling $a = g^2/(4\pi)^2$, which is obviously a convenient choice, we follow [**DeKaTh96, DeKaTh97**] and other authors whose results we shall come to in due course. The meromorphic function

$$F(\rho) := \int \frac{d^4 k_E}{2\pi^2} \frac{(k_E^2)^{-\rho}}{k_E^2 (k_E - q_E)^2} \bigg|_{q_E^2 = 1} = \frac{1}{2\rho(1-\rho)} \tag{5.2.10}$$

is referred to as the *Mellin transform* of the skeleton graph  [**Krei06, KrY06, Y11**].

5.2.2. Rainbow approximation. We will see now how this function naturally arises in the rainbow DSE

(5.2.11)



which, in its analytic form reads

$$-i\Sigma_{RB}(q) = (-ig)^2 \int \frac{d^4 k}{(2\pi)^4} \frac{i}{\not{k} + i\epsilon} \left[1 - i\Sigma_{RB}(k) \frac{i}{\not{k} + i\epsilon} \right] \frac{i}{(q-k)^2 + i\epsilon}. \tag{5.2.12}$$

Since we can write the self-energy in terms of a single form factor, ie $-i\Sigma_{RB}(q) = \not{q} A(q^2)$, this equation reduces to

$$A(-q_E^2) = ia \int \frac{d^4 k_E}{2\pi^2} \frac{1}{k_E^2 (q_E - k_E)^2} \left[1 + iA(-q_E^2) \right] \tag{5.2.13}$$

in Euclidean form, ie after Wick rotation. Renormalised in momentum scheme, this equation morphs into

$$A_R(q_E^2, \mu^2) = ia \int \frac{d^4 k_E}{2\pi^2} \left\{ \frac{1}{k_E^2 (q_E - k_E)^2} - \frac{1}{k_E^2 (\widetilde{q}_E - k_E)^2} \right\} \left[1 + iA_R(k_E^2, \mu^2) \right] \tag{5.2.14}$$

in which $A_R(q_E^2, \mu^2) := A(-q_E^2) - A(-\mu^2)$ is the renormalised cousin of the form factor and $\widetilde{q}_E$ is the Euclidean reference momentum with reference (renormalisation) scale $\mu > 0$, ie $\widetilde{q}_E^2 = \mu^2$. The Mellin transform emerges if we try a *scaling ansatz*

$$G(a, \ln(q_E^2/\mu^2)) := 1 + iA_R(q_E^2, \mu^2) = \left(\frac{q_E^2}{\mu^2} \right)^{-\gamma(a)} \tag{5.2.15}$$

for what we call the *Green's function* of the rainbow approximation. $\gamma(a)$ is a yet unknown function of the coupling a about which we aim to get some information. Because of its position in the log-expansion

$$G(a, \ln(q_E^2/\mu^2)) = \sum_{n\geq 0} \frac{(-1)^n}{n!} \gamma(a)^n \ln^n(q_E^2/\mu^2) \tag{5.2.16}$$

it makes sense to call it the *anomalous dimension (of the Yukawa fermion)*. When we insert this ansatz into (5.2.14), we get

$$\begin{aligned} \left(\frac{q_E^2}{\mu^2}\right)^{-\gamma(a)} &= 1 - a \int \frac{d^4k_E}{2\pi^2} \left\{ \frac{1}{k_E^2(q_E - k_E)^2} - \frac{1}{k_E^2(\widetilde{q}_E - k_E)^2} \right\} \left(\frac{k_E^2}{\mu^2}\right)^{-\gamma(a)} \\ &= 1 - a \left\{ \left(\frac{q_E^2}{\mu^2}\right)^{-\gamma(a)} - 1 \right\} F(\gamma(a)) \end{aligned} \tag{5.2.17}$$

which entails

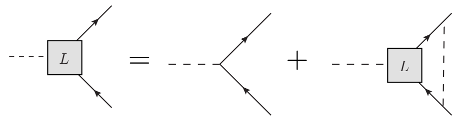
$$1 = -aF(\gamma(a)) = \frac{a}{2\gamma(a)(\gamma(a)-1)}. \tag{5.2.18}$$

If a DSE can be solved with a scaling ansatz, then an implicit equation like this one is what one should aim for [**Krei06**]. In this simple (but nontrivial) rainbow case, we have a luxurious situation of a quadratic equation (5.2.18) for the anomalous dimension whose solution is given by

$$\gamma^{\pm}(a) = \frac{1 \pm \sqrt{1+2a}}{2}. \tag{5.2.19}$$

We select $\gamma(a) = \gamma^-(a)$ on the grounds that it is 'more physical' due to satisfying the condition $\gamma(0) = 0$; 'more physical' to the extent that a rainbow approximation can be physical[2].

5.2.3. Ladder approximation. The same method can be applied to the ladder DSE [**Krei06**],

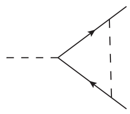
(5.2.20) 

,

ie explicitly

$$-ig\Gamma(q) = -ig + (-ig)^3 \int \frac{d^4k}{(2\pi)^4} \frac{i}{\not{k} + i\epsilon} \Gamma(k) \frac{i}{\not{k} + i\epsilon} \frac{i}{(q-k)^2 + i\epsilon}, \tag{5.2.21}$$

where the Mellin transform of the skeleton

(5.2.22) $u =$  $= (-ig)^3 \int \frac{d^4k}{(2\pi)^4} \frac{i}{\not{k} + i\epsilon} \frac{i}{\not{k} + i\epsilon} \frac{i}{(q-k)^2 + i\epsilon},$

is given by

$$F_u(\rho) = \frac{1}{\rho(1-\rho)} \tag{5.2.23}$$

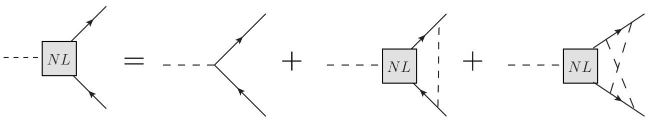
and the scaling ansatz $\Gamma(q_E) = (q_E^2/\mu^2)^{-\gamma_u(a)}$ leads again to $1 = -aF_u(\gamma_u(a))$. The (physical) solution for the anomalous dimension $\gamma_u(a)$ is similiar to that of the rainbow:

$$\gamma_u(a) = \frac{1 - \sqrt{1+4a}}{2}, \tag{5.2.24}$$

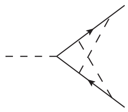
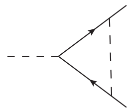
[2]It is not very physical, it lacks an important feature: it fails to have a divergent Taylor series.

again an algebraic function of the coupling a with a convergent Taylor series.

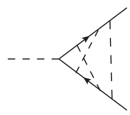
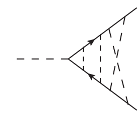
5.2.4. Next-to-ladder approximation. The above strategy has been successfully applied to more interesting cases, albeit with the drawback that the implicit equation for the anomalous dimension cannot be solved analytically. The so-called *next-to-ladder approximation* for which the DSE takes the form

(5.2.25) 

has been tackled in [**BiKreiW07**] for zero momentum transfer, which means that the external boson has vanishing momentum. This equation has an extra primitive 'skeleton' graph given by

(5.2.26) $v =$  in addition to $u =$ 

which introduces graphs like

(5.2.27) $vu =$  and $uuv =$ 

into the game. To be more precise, the resulting ladders have two types of rungs, denoted by the letters u and v. This notation makes it obvious that one can express all graphs in the next-to-ladder series as words comprised of the two letters u and v. In combinatorial notation, the next-to-ladder DSE (5.2.25) is written as

$$X_{NL}(a) = \mathbb{I} + aB_+^u(X_{NL}(a)) + a^2 B_+^v(X_{NL}(a)), \tag{5.2.28}$$

where the two linear insertion operators are defined by $B_+^v(w) = wv$ and $B_+^u(w) = wu$ for any word w made up of the letters u and v. The empty word is $\mathbb{I}$ for which $B_+^u(\mathbb{I}) = u$ and $B_+^v(\mathbb{I}) = v$.

We obtain the combinatorial *ladder DSE* if we drop the second insertion operator in (5.2.28), and restrict ourselves to rung type u, ie

$$X_L(a) = \mathbb{I} + aB_+^u(X_L(a)). \tag{5.2.29}$$

The combinatorial solution of this equation is simply $X_L(a) = \mathbb{I} + \sum_{k\geq 1} a^k u^k = (\mathbb{I} - au)^{-1}$, whereas for the next-to-ladder case we find

$$X_{NL}(a) = \mathbb{I} + au + a^2(uu + v) + a^3(uuu + uv + vu) + ... \tag{5.2.30}$$

by using the ansatz $X_{NL}(a) = \mathbb{I} + \sum_{k\geq 1} a^k x_k$ and plugging it into (5.2.28). The method of Mellin transforms leads to the implicit equation

$$1 = -aF_u(\gamma_G(a)) + a^2 F_v(\gamma_G(a)) \ , \tag{5.2.31}$$

where the two functions F_u and F_v are the Mellin transforms of the skeleton graphs u and v, respectively. As we have already alluded to, (5.2.31) can only be solved numercially. For details, the reader is referred to [**BiKreiW07**].

5.2.5. General linear DSEs. Note that the combinatorial DSE of the rainbow approximation has the same form as that of the ladder approximation, ie the form of (5.2.29). These and the next-to-ladder DSE (5.2.28) have one important feature in common: they fall all into the category of so-called *linear DSE*. The motive for this denomination is that in these three cases, the combinatorial series $X(a)$ does *not* appear in higher powers, as is the case in

$$X(a) = \mathbb{I} + aB_+(X(a)^2) \tag{5.2.32}$$

which an example of a nonlinear DSE that emerges as an approximation in the massless supersymmetric Wess-Zumino model [**BeCla14**]. However, as a consequence of being nonlinear, the insertion operators in a nonlinear DSE get confronted with nontrivial products of 1PI graphs in their arguments, while this never happens in the linear case. Nontrivial products of graphs lead to fundamentally different solutions.

The most general form of linear DSEs for amplitudes $X_1(a), ..., X_N(a)$ having been investigated so far are given by the system

$$X_k(a) = \mathbb{I} + a \sum_{l=1}^{N} \sum_{j=1}^{N} M_{kj}^{(l)} B_+^l(X_j(a)) \qquad (k = 1, ..., N), \tag{5.2.33}$$

where $M^{(l)}$ is a (real) $N \times N$ matrix associated to the primitive 1-loop skeleton l. My colleague Henry Kißler has shown in master's thesis [**Ki12**] that if the matrix of Mellin transforms resulting from (5.2.33) is diagonalisable, then the solution of the analytic version of (5.2.33) is a linear combination of scaling factors:

$$G_k(a, L) = \sum_{j=1}^{N} c_{kj}(\{\gamma(a)\}) e^{-\gamma_j(a)L}, \qquad L := \log(-q^2/\mu^2), \tag{5.2.34}$$

where the expressions $c_{kj}(\{\gamma(a)\})$ are algebraic functions of the corresponding collection of anomalous dimensions $\{\gamma(a)\} = \{\gamma_j(a) : j = 1, ..., N\}$, one for each amplitude. Although Kißler's proof employs Feynman rules of a toy model, the involved steps leave no doubt that the assertion holds in general. This author's work shows nicely the virtue of the combinatorial approach and the notation being used: a translation of (5.2.33) into a system of blob-diagram equations is impossible as there is no way to draw a 'general skeleton' let alone incorporate blobs into it.

We mention for completeness that, as regards Yukawa theory, the author has furthermore solved for $N = 2$ the Yukawa approximation

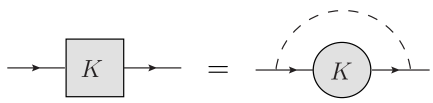
$$\begin{aligned} X &= \mathbb{I} - \alpha B_+ (X + X) \\ X &= \mathbb{I} - \alpha B_+ (X) \end{aligned} \tag{5.2.35}$$

explicitly (see [**Ki12**], Chapter 5).

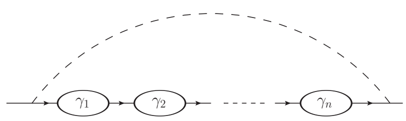
5.2.6. Nonlinear DSE: Kilroy approximation. We shall now discuss a nonlinear example: the so-called *Kilroy Dyson-Schwinger equation*[3]

$$X(a) = \mathbb{I} - aB_+(1/X(a)) \qquad \text{'Kilroy DSE'}, \tag{5.2.36}$$

which, if applied to the self-energy of Yukawa fermions reads

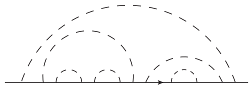
(5.2.37) 

in blob-diagrammatical form, where the round blob on the right represents the full propagator. This DSE describes an approximation for the self-energy in which graphs of the form

(5.2.38) $\Gamma =$  $n \geq 0$

[3]The reader is encouraged to find out the motivation for this name (see internet) and why it is not entirely appropriate.

emerge. In it, any subgraph γ_j is either a rainbow graph or a graph one obtains by any sequence of consecutive but non-overlapping corrections on the fermion line of the self-energy 1-loop graph. This means that any chainings and nestings of rainbow subgraphs are involved, eg

(5.2.39)  (Kilroy graph).

In its analytical form, the Kilroy DSE reads for Yukawa theory

$$-i\Sigma_K(q) = (-ig)^2 \int \frac{d^4k}{(2\pi)^4} \frac{i}{\not k + i\epsilon} \left[1 + i\Sigma_K(k) \frac{i}{\not k + i\epsilon}\right]^{-1} \frac{i}{(q-k)^2 + i\epsilon} \,. \tag{5.2.40}$$

Not surprisingly, this equation cannot be solved by a simple scaling ansatz like in the case of the above linear DSEs. Broadhurst and Kreimer have investigated the Kilroy DSE for both Yukawa and $(\varphi^3)_6$ theory in [**BroK01**]. The upshot is that the anomalous dimension $\gamma_K(a)$ of the Kilroy approximation (5.2.40) satisfies the implicit equation

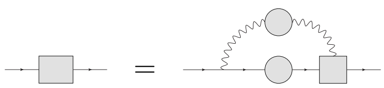
$$\sqrt{\frac{a}{\pi}} e^{-Z(a)} = 1 + \operatorname{erf}(Z(a)), \tag{5.2.41}$$

where $Z(a) = (\gamma_K(a) - 1)/\sqrt{a}$ and $\operatorname{erf}(x)$ is the famous error function. Broadhurst solved this implicit equation[4] numercially for $\gamma_K(a)$ by an algorithm of the Newton-Raphson type to very high loop orders ($\sim$ 500, for details see [**BroK01**]). In particular, it turns out that the anomalous dimension of the Kilroy approximation has a divergent perturbation series! This sets it apart from the results of the rainbow and the two ladder approximations and makes it 'more physical'.

We shall see that (5.2.41) can also be derived through an ordinary differential equation in Section 5.6 where we revisit the Kilroy DSE [**KrY06**].

5.3. Dyson-Schwinger equations of QED

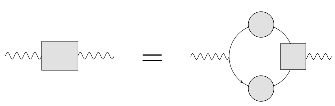
5.3.1. Integral equations and blob diagrams. In the language of blob diagrams, the DSEs of quantum electrodynamics may be formulated as follows [**BjoDre65**]. For the fermion self-energy, one has

(5.3.1)  (fermion self-energy)

which encodes the analytic equation

$$-i\Sigma(q) = (-ie)^2 \int \frac{d^4k}{(2\pi)^4} \gamma^\mu iS(k)\Gamma^\nu(k, q-k) iD_{\mu\nu}(q-k), \tag{5.3.2}$$

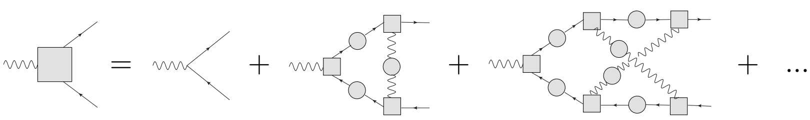
and

(5.3.3)  (photon self-energy)

represents the DSE of the photon's self-energy:

$$q^2\pi(q^2) = i\frac{(-ie)^2}{3} \int \frac{d^4k}{(2\pi)^4} \operatorname{tr}[\gamma^\mu iS(k)\Gamma_\mu(k, q-k) iS(q-k)]. \tag{5.3.4}$$

The DSE for the vertex function can be formally written as a skeleton series

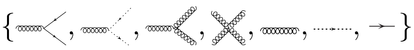
(5.3.5)  (vertex function)

[4]The authors used different conventions, for details see Section 5.6.

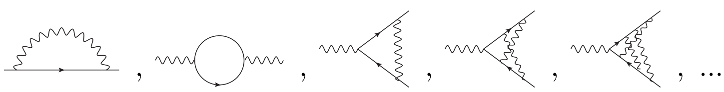
for which we abstain from writing out the analytic form. The problem with this equation is that it represents an integral equation with an infinite number of integral operators on the rhs, one for each skeleton, and no one knows whether a solution exists.

This corresponds to the question of the existence of QED, apart from the Landau pole issue. However, we can rightfully expect there to be a solution for any finite number of skeletons. We take the optimistic view that renormalised QED exists: given that there is a solution for each set of DSE skeletons taken into account, we expect this so-defined sequence of solutions to converge and yield something sensible. What nourishes this view is that regarding the combinatorial DSE, to be discussed next, it is clear that by construction, a solution *does* exist in the set $\mathcal{H}[[\alpha]]$.

5.3.2. Combinatorial DSE. The overall divergent amplitudes of a renormalisable theory can be conveniently characterised by the corresponding vertex and propagator types, for example, in QED, they are given by $\mathcal{R}_{\mathrm{QED}} = \{$  $\}$ while the case of QCD is a bit richer, namely

(5.3.6) $\mathcal{R}_{\mathrm{QCD}} = \{$  $\}$.

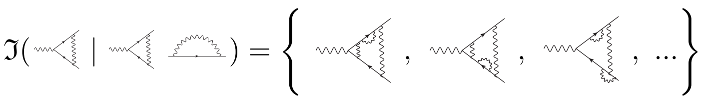
These sets are *finite*, as we would otherwise not be dealing with a renormalisable field theory. We call such sets the *residue set* of a theory[5]. For every divergent amplitude $r \in \mathcal{R}$, there is a set of primitive diagrams, where in the case of QED, the primitive diagrams are given by

(5.3.7) 

which serve as skeletons in the DSE (5.3.1) to (5.3.5). For the combinatorial description, we define in QED a linear insertion operator B_+^p for each primitive skeleton graph p by

$$B_+^p(G) := \sum_{\Gamma \in \mathfrak{I}(p|G)} \frac{\Gamma}{|\mathfrak{I}(p|G)|}\ , \tag{5.3.8}$$

where G is a product of 1PI graphs and $\mathfrak{I}(p|G)$ is the set of all graphs that one can obtain from inserting the graph G into p. As an example of the set $\mathfrak{I}(p|G)$, consider

(5.3.9) 

which has 6 elements (exercise for the reader). Notice that both graphs are inserted simultaneously into the skeleton. A general and highly non-trivial definition of insertion operators for DSEs, can for example be found in [**Krei06a**].

With these operators, we can now write the DSE in the form of a system of combinatorial equations:

$$X^r(\alpha) = \mathbb{I} + \mathrm{sgn}(s_r) \sum_{\mathrm{res}(p)=r} \alpha^{|p|} B_+^p(Q(\alpha)^{|p|} X^r(\alpha))\ , \qquad r \in \mathcal{R}. \tag{5.3.10}$$

This form of the DSE is very general and makes sense for any renormalisable quantum field theory with only one coupling parameter (eg QED, QCD) [**Krei06a, Y11**]. The notation means the following. First, the sum (5.3.10) ranges over all primitive skeletons with residue r. Second, α is the coupling parameter[6] and Q is the *invariant charge*, a combinatorial series defined by the product

$$Q = \prod_{r \in \mathcal{R}} (X^r)^{s_r} \tag{5.3.11}$$

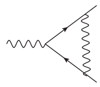
[5]The reader be reminded at this point of the concept of the residue of a graph, Definition 4.5 Section 4.3.

[6]α is the fine-structure constant in the case of QED.

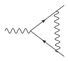
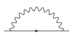
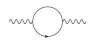
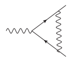
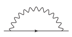
in which the numbers $s_r \in \mathbb{Z}$ depend on the residue. For QED, as we have already seen in (4.3.18) in the previous section, this takes the concrete form

$$Q = \frac{(X^{\prec})^2}{X^{\sim}(X^{\to})^2} \; . \tag{5.3.12}$$

The exponents s_r depend on the number of insertion places the $|p|$-loop primitive p offers for a graph of residue r. For example,

(5.3.13) $p =$ 

has three insertion places for vertex corrections, two for fermion propagator corrections and one for photon propagator corrections on offer. The product series

(5.3.14) $Q(\alpha)X^{\prec}(\alpha) = \mathbb{I} + \alpha \Big(3$  $+ 2$  $+$  $\Big) + \alpha^2 \Big(6$   $+ \; ... \Big) + \; ...$

presents exactly the right graphs destined for insertion into the 1-loop vertex skeleton p, which is why this expression appears inside the argument of the insertion operator B_+^p in (5.3.10). The coefficients indicate how many possibilities there are to insert the corresponding graphs into the skeleton which is precisely the cardinality of the set $\mathfrak{I}(p|G)$, where we recall that G is the graph to be inserted into the skeleton p. With these ingredients, we can write the combinatorial form of the DSE in QED as

$$\begin{aligned} X^{\to} &= \mathbb{I} - \alpha B_+^{\frown}(QX^{\to}) \; , \qquad X^{\sim} = \mathbb{I} - \alpha B_+^{\circ}(QX^{\sim}), \\ X^{\prec} &= \mathbb{I} + \sum_{j \geq 1} \alpha^j B_+^{p_j}(Q^j X^{\prec}) = \mathbb{I} + \alpha B_+^{\lhd}(QX^{\prec}) + ... \end{aligned} \tag{5.3.15}$$

in which $p_1, p_2, ...$ are the primitive vertex skeletons, the index being the loop number, ie $|p_j| = j$. One can pass over to the analytic DSE by applying a Feynman character to both sides, preferably the renormalised one, where the insertion operator is intertwined into an integral operator

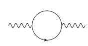
$$(\chi_R \circ B_+^{\gamma})(G) = \int \mathscr{R}[\omega_\gamma](\chi_R(G)). \tag{5.3.16}$$

$\int \mathscr{R}[\omega_\gamma](...)$ is the renormalised integral operator that corresponds to the skeleton γ and G is a graph. Because the insertion operator's superscript γ is primitive, the integral kernel needs only one subtraction, ie $\mathscr{R}[\omega_\gamma] = \omega_\gamma + \mathscr{C}[\omega_\gamma]$.

5.3.3. DSEs in the quotient Hopf algebra of QED. In Section 4.3 we have introduced the Hopf ideal $\mathcal{I}$ generated by the Ward elements (4.3.13) and the resulting quotient Hopf algebra

$$\mathcal{H}_\sim = \mathcal{H}/\mathcal{I} \tag{5.3.17}$$

in which the invariant charge Q takes the simplified form $Q_\sim = \mathbb{I}/X^{\sim}$. The problem is, the DSEs in (5.3.15) cease to make sense in $\mathcal{H}_\sim$ and 'degenerate' into an approximation: the photon series terminates after the 1-loop contribution

(5.3.18) $\mathbb{I} - \alpha B_+^{\circ}(Q_\sim X^{\sim}) = \mathbb{I} - \alpha B_+^{\circ}(\mathbb{I}) = \mathbb{I} - \alpha$  ,

the fermion self-energy receives no vertex corrections and the vertex no fermion line corrections. The QCD case is different in this respect. Passing over to the quotient Hopf algebra there, is, in fact, necessary to formulate the combinatorial DSEs of QCD in the form (5.3.10) in the first place, see [**Krei06a**] for an exposition of these equations[7].

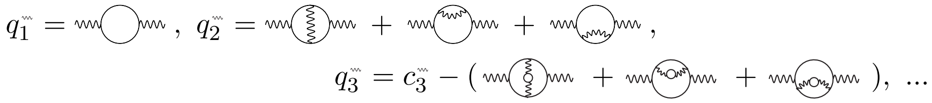
[7]In QED, the rub lies in having to equate $\mathbb{I}$ to the quotient of two series on account of the Ward identity (4.3.24), while at the same time $\mathbb{I}$ is needed for generating the first-order term. A technical blemish.

The remedy is to replace the superscript graphs by the combinatorial coefficients of *quenched QED*[8] which for the quenched photon series

$$X^{\sim} = \mathbb{I} - \sum_{j\geq 1} q_j^{\sim} \alpha^j \tag{5.3.19}$$

are

(5.3.20)



and so on. For the photon's DSE, we replace the insertion operator $B_+^{\sim\circ\sim}$ by

$$B_+^{q_j^{\sim}} := \sum_s B_+^{\gamma_{j,s}} \tag{5.3.21}$$

where $q_j^{\sim} = \sum_s \gamma_{j,s}$ is the j-th contribution to the quenched photon propagator, a sum of 1PI photon propagator graphs $\gamma_{j,1}, \gamma_{j,2}, \ldots$ with no subgraph having a closed fermion loop. This new insertion operator enables us to write the photon's DSE as an equation that decouples from the other two (which also have to be modified accordingly):

$$X^{\sim} = \mathbb{I} - \sum_{j\geq 1} \alpha^j B_+^{q_j^{\sim}}(Q_{\sim}^j X^{\sim}) = \mathbb{I} - \sum_{j\geq 1} \alpha^j B_+^{q_j^{\sim}}((X^{\sim})^{1-j}) \, . \tag{5.3.22}$$

If we apply the renormalised Feynman character, this expression becomes

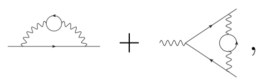
$$\chi_R(X^{\sim}(\alpha)) = 1 - \sum_{j\geq 1} \alpha^j (\chi_R \circ B_+^{q_j^{\sim}})((X^{\sim})^{1-j}) = 1 - \sum_{j\geq 1} \alpha^j \int \mathscr{R}[\omega_{q_j^{\sim}}](\chi_R(X^{\sim})^{1-j}). \tag{5.3.23}$$

This is the form factor of the photon's self-energy. The notation on the rhs means that in the j-th term, this function is found in $(j-1)$ different insertion places. Notice that on account of the Ward identity, the coefficients of the quenched series in (5.3.20) are all primitive, if we modify the Hopf ideal $\mathcal{I}$ as follows. Every Ward element[9]

$$w_j = c_j^{\sim\!<} + c_j^{-\!\!-} \tag{5.3.24}$$

can be split into two components $w_j = \widetilde{w}_j + \overline{w}_j$, where $\overline{w}_j$ is quenched, ie it has no internal fermion loops. Take w_2 for example. If we subtract the quenched contribution from w_2, we get the non-quenched part,

$$\widetilde{w}_2 = w_2 - \overline{w}_2 = \tag{5.3.25}$$



while $\widetilde{w}_1 = w_1 - \overline{w}_1 = 0$, because there is no internal fermion loop at 1-loop level.

Since $S_R^{\chi}(\widetilde{w}_j) = 0$ for any j, one can mod out both components of w_j individually. Then follows that we need just one subtraction for all new Dyson-Schwinger skeletons in the quotient Hopf algebra of QED, ie $\mathscr{R}[\omega_{q_j^{\sim}}] = \omega_{q_j^{\sim}} + \mathscr{C}[\omega_{q_j^{\sim}}]$ for all j.

5.4. Renormalisation group recursion and Callan-Symanzik equation

We shall in this section present the interesting result that the coproduct formula

$$\Delta(X^r) = X^r \otimes \mathbb{I} + \mathrm{sgn}(s_r) \sum_{n\geq 1} X^r Q^n \otimes c_n^r \alpha^n, \tag{5.4.1}$$

[8]In quenched QED, there is no subgraph with a closed fermion loop, all photon propagator subgraphs are replaced by a bare photon propagator.

[9]See Section 4.3.

from Proposition 4.6 provides the basis for a combinatorial derivation of the Callan-Symanzik equation. Moreover, it implies a recursion formula for what we will refer to as the *log-coefficient* or *RG functions*.

5.4.1. Infinitesimal characters as Lie generators. Except for some pathological cases with subdivergences leading to nonsensical cographs[10], it is well-known that in momentum scheme the renormalised form factor of any (divergent) single-scale 1PI graph Γ with external momentum $q \in \mathbb{M}$ evaluates to a polynomial of the form

$$\chi_R(\Gamma) = \sum_{j=1}^{N} \sigma_j(\Gamma) L^j, \tag{5.4.2}$$

where $\sigma_j(\Gamma) \in \mathbb{R}$ are the coefficients, $L = \ln(-q^2/\mu^2)$ is the external momentum parameter and $\mu > 0$ the renormalisation reference scale. We remind the reader that we do not include the coupling parameter in the definition of the Feynman characters because we apply them in the combinatorial formalism to formal power series in this very coupling. The degree N of the polynomial with $1 \leq N \leq |\Gamma|$ depends on the number of subdivergences: $N = \#(\text{subdivergences})+1$.

We can view the family $\{\sigma_j\}$ as linear maps on the linear span of all Feynman graphs of the theory in question which evaluate a graph to the corresponding coefficient of the polynomial in L. This implies $\sigma_j(\Gamma) = 0$ if $j > |\Gamma|$, ie whenever the index exceeds the loop number.

We will now see that a renormalised Feynman character χ_R is generated by a linear map $\sigma\colon \mathcal{H} \to \mathbb{R}$ which is closely related to the maps σ_j. Consider the next

DEFINITION 5.1 (Infinitesimal characters). *A linear map $\sigma : \mathcal{H} \to \mathbb{R}$ such that*

$$\sigma(xy) = e(x)\sigma(y) + \sigma(x)e(y) \qquad \textit{for all} \quad x, y \in \mathcal{H} \tag{5.4.3}$$

is called infinitesimal character*.*

Note that this definition implies in particular that $\sigma(\mathbb{I}) = 0$ and that all nontrivial products of graphs are also mapped to zero. This is caused by $\mathrm{e}(G) = 0$ for $G \neq \mathbb{I}$. We denote the linear space of such maps by $\mathrm{ch}(\mathcal{H}, \mathbb{R})$. This space is not closed with respect to the $\star$-convolution:

$$(\sigma \star \tau)(G_1 G_2) = (\sigma \otimes \tau)(G_1 \otimes G_2 + G_2 \otimes G_1 + ...) = \sigma(G_1)\tau(G_2) + \sigma(G_2)\tau(G_1) \tag{5.4.4}$$

which does not vanish in general if $G_1, G_2 \neq \mathbb{I}$ (the remainder '...' consists of nontrivial products on at least on side of the tensor product and is therefore mapped to zero).

Therefore, $\mathrm{ch}(\mathcal{H}, \mathbb{R})$ is neither a convolution group nor algebra. Instead, it is closed under the Lie bracket

$$[\sigma, \tau]_\star := \sigma \star \tau - \tau \star \sigma \ , \tag{5.4.5}$$

and because the Jacobi identity is satisfied by the $\star$-convolution's associativity, $(\mathrm{ch}(\mathcal{H}, \mathbb{R}), [\cdot, \cdot]_\star)$ is in fact a Lie algebra. The next assertion tells us that these maps generate Hopf algebra characters $\chi\colon \mathcal{H} \to \mathbb{R}$ which is why they are sometimes referred to as *Lie generators*.

PROPOSITION 5.2. $\mathrm{Ch}(\mathcal{H}, \mathbb{R}) = \exp_\star(\mathrm{ch}(\mathcal{H}, \mathbb{R}))$*, ie if* $\sigma \in \mathrm{ch}(\mathcal{H}, \mathbb{R})$ *then*

$$\chi = \exp_\star \sigma = \sum_{n=0}^{\infty} \frac{\sigma^{\star n}}{n!} \in \mathrm{Ch}(\mathcal{H}, \mathbb{R}) \tag{5.4.6}$$

and conversely, for each $\chi \in \mathrm{Ch}(\mathcal{H}, \mathbb{R})$*, there is a* $\sigma \in \mathrm{ch}(\mathcal{H}, \mathbb{R})$ *such that (5.4.6) holds.*

[10] In a renormalisable theory like QED these fellows always find pathological partners on each loop order which render each other harmless. Otherwise, QED would not be renormalisable.

Proof. By induction, one finds that $\sigma^{\star n}(xy) = \sum_{j=0}^{n} \binom{n}{j} \sigma^{\star j}(x)\sigma^{\star n-j}(y)$. Then multiplicativity follows immediately

$$\exp_\star \sigma(xy) = \sum_{n\geq 0} \frac{1}{n!} \sum_{j=0}^{n} \binom{n}{j} \sigma^{\star j}(x)\sigma^{\star n-j}(y) = \exp_\star \sigma(x) \exp_\star \sigma(y) \ . \tag{5.4.7}$$

Linearity is inherited from both product and coproduct. Hence $\exp_\star \sigma$ is a character. To see that

$$\sigma := \log_\star \chi = -\sum_{n\geq 1} \frac{1}{n} (\mathrm{e} - \chi)^{\star n} \tag{5.4.8}$$

is an infinitesimal character, first note that

$$(\mathrm{e} - \chi)^{\star n}(xy) = (\mathrm{e} \otimes \mathrm{e} - \chi \otimes \chi)^{\star n}(x \otimes y) \tag{5.4.9}$$

and

$$(\mathrm{e} \otimes \mathrm{e} - \chi \otimes \mathrm{e})^{\star n} = ([\mathrm{e} - \chi] \otimes \mathrm{e})^{\star n} = (\mathrm{e} - \chi)^{\star n} \otimes \mathrm{e}. \tag{5.4.10}$$

Therefore

$$\begin{aligned} \log_\star \chi(xy) &= \log_\star(\chi \otimes \chi)(x \otimes y) = \log_\star(\underbrace{[\chi \otimes \mathrm{e}] \star [\mathrm{e} \otimes \chi]}_{=(\chi\star \mathrm{e})\otimes(\mathrm{e}\star\chi)=\chi\otimes\chi})(x \otimes y) \\ &= (\log_\star[\chi \otimes \mathrm{e}] + \log_\star[\mathrm{e} \otimes \chi])(x \otimes y) \\ &= (\log_\star \chi \otimes \mathrm{e} + \mathrm{e} \otimes \log_\star \chi)(x \otimes y) = \log_\star \chi(x)\mathrm{e}(y) + \mathrm{e}(x) \log_\star \chi(y) \end{aligned} \tag{5.4.11}$$

□

Notice that this result is rather general: we could have formulated it for any target set $\mathcal{A}$ of the characters as long as it is a commutative algebra. Because renormalised Feynman characters map single-scale Feynman graphs to elements in the polynomial algebra $\mathcal{A} = \mathbb{R}[L]$, it is this very target algebra that will interest us in the following. It turns out that the elements of the character group $\mathrm{Ch}(\mathcal{H}, \mathbb{R}[L])$ are all of the form

$$\chi = \exp_\star(L\sigma) = \sum_{n\geq 0} \frac{L^n}{n!} \sigma^{\star n}, \tag{5.4.12}$$

where $\sigma \in \mathrm{ch}(\mathcal{H}, \mathbb{R})$ is a Lie generator. This result by [**KreiPa12**] is worth a proposition.

Proposition 5.3. *Let $\chi \in \mathrm{Ch}(\mathcal{H}, \mathbb{R}[L])$ be a coalgebra homomorphism*[11]*, ie $(\chi \otimes \chi)\Delta = \Delta\chi$. Then $\chi = \exp_\star(L\partial_0\chi)$, where $\partial_0 \colon \mathbb{R}[L] \to \mathbb{R}[L]$ is the linear map given by*

$$\partial_0(L^n) = \delta_{n,1} = \begin{cases} 1 & n = 1 \\ 0 & \mathit{else} \end{cases} \tag{5.4.13}$$

and $\partial_0\chi := \partial_0 \circ \chi \in \mathrm{ch}(\mathcal{H}, \mathbb{R})$ is an infinitesimal character.

Proof. First note that the evaluation maps $\mathrm{ev}_a \colon \mathbb{R}[L] \to \mathbb{R}$, $p(L) \mapsto p(a)$, constitute the character group $\mathrm{Ch}(\mathbb{R}[L], \mathbb{R})$, where we recall that $\mathbb{R}[L]$ is a Hopf algebra.

If η is any such character, then it is uniquely determined by $\lambda := \eta(L)$, in fact, it is the character ev_λ, because of

$$\eta(p(L)) = p(\eta(L)) = p(\lambda) = \mathrm{ev}_\lambda(p(L)). \tag{5.4.14}$$

The reader may check that $\mathrm{ev}_\lambda = \exp_\star(\lambda\partial_0)$ by using $\partial_0^{\star k}(L^n) = n!\delta_{k,n}$ (easily verified by induction). Finally, by the coalgebra morphism property,

$$\mathrm{ev}_\lambda \circ \log_\star \chi = \log_\star(\mathrm{ev}_\lambda \circ \chi) = \log_\star(\mathrm{ev}_\lambda) \circ \chi = \lambda\partial_0 \circ \chi = \lambda\partial_0\chi. \tag{5.4.15}$$

□

[11] The target algebra $\mathcal{A} = \mathbb{R}[L]$ is a coalgebra, see Appendix A.2.

5.4.2. Callan-Symanzik equation. This latter result has interesting consequences. Assume now that $\chi \in \mathrm{Ch}(\mathcal{H}, \mathbb{R}[L])$ is a renormalised Feynman character for single-scale graphs. It is then of the form (5.4.12) and the first derivative with respect to the momentum parameter yields $\partial_L \chi = \sigma \star \chi$ which, if we apply it to a combinatorial series X^r entails

$$\partial_L \chi(X^r(\alpha)) = (\sigma \star \chi)(X^r(\alpha)) = (\sigma \otimes \chi)\Delta(X^r(\alpha)) \tag{5.4.16}$$

and using the coproduct formula (5.4.1), we find [**Y11**]

$$\partial_L \chi(X^r(\alpha)) = \sigma(X^r(\alpha)) + \mathrm{sgn}(s_r) \sum_{n \geq 1} \sigma(X^r Q^n) \chi(c_n^r) \alpha^n. \tag{5.4.17}$$

This equation is in fact a Callan-Symanzik equation for the Green's function $G^r(\alpha, L) := \chi(X^r(\alpha))$, as the next assertion shows [**Y11**].

PROPOSITION 5.4 (Callan-Symanzik equation). *The Green's function $G^r(\alpha, L)$ satisfies the differential equation*

$$\left[-\frac{\partial}{\partial L} + \beta(\alpha) \frac{\partial}{\partial \alpha} + \mathrm{sgn}(s_r) \gamma^r(\alpha) \right] G^r(\alpha, L) = 0 \tag{5.4.18}$$

where $\gamma^r(\alpha) := \mathrm{sgn}(s_r)\sigma(X^r(\alpha))$ is the anomalous dimension and $\beta(\alpha) := \alpha\sigma(Q(\alpha)) = \alpha \sum_{t \in \mathcal{R}} |s_t| \gamma^t(\alpha)$ the β-function of the corresponding theory with amplitude set $\mathcal{R}$.

PROOF. This equation follows directly from (5.4.17). First, note that

$$\sigma(X^r Q^n) = \sigma(X^r) + n\sigma(Q), \tag{5.4.19}$$

which is implied by $\sigma(\mathbb{I}) = 0$ and (5.4.3), in particular, if we decompose $Q = \mathbb{I} + \overline{Q}$, we get

$$\sigma(Q^n) = \sigma((\mathbb{I} + \overline{Q})^n) = \sigma(\mathbb{I} + n\overline{Q}) = n\sigma(\overline{Q}) = n\sigma(Q) \tag{5.4.20}$$

because Q starts with $\mathbb{I}$ and the remainder, namely $\overline{Q}$ has only nontrivial coefficients from $\mathcal{H}$ such that powers of $\overline{Q}$ have only nontrivial products of graphs which lie in the kernel of the infinitesimal character σ. For the charge Q, we have

$$Q = \prod_{t \in \mathcal{R}} (X^t)^{s_t} = \prod_{t \in \mathcal{R}} (\mathbb{I} + \overline{X}^t)^{s_t} = \mathbb{I} + \sum_{t \in \mathcal{R}} s_t \overline{X}^t + \ldots\,, \tag{5.4.21}$$

where the rest has only nontrivial products of graphs from higher powers of $\overline{X}^t$. Then follows

$$\sigma(Q(\alpha)) = \sum_{t \in \mathcal{R}} s_t \sigma(\overline{X}^t) = \sum_{t \in \mathcal{R}} s_t \sigma(X^t) = \sum_{t \in \mathcal{R}} s_t \mathrm{sgn}(s_r)^2 \sigma(X^t) = \sum_{t \in \mathcal{R}} |s_t| \gamma^t(\alpha) \tag{5.4.22}$$

If we plug all this into (5.4.19) the resulting expression into (5.4.17), we get the Callan-Symanzik equation (exercise for the reader). □

In QED, we have the luxury of the Ward identity, which means for the β-function that it depends only on the anomalous dimension of the photon:

$$\beta(\alpha) = \alpha \sum_{t \in \mathcal{R}} |s_t| \gamma^t(\alpha) = \alpha[\underbrace{2\gamma^{\rightarrow}(\alpha) + 2\gamma^{\sim\!<}(\alpha)}_{=0} + \gamma^{\sim}(\alpha)] = \alpha\gamma^{\sim}(\alpha). \tag{5.4.23}$$

Linear DSEs like the rainbow and the two ladder approximations from Section 5.2 have a trivial charge $Q = \mathbb{I}$ and hence $\beta(a) = 0$. The Callan-Symanzik equation then trivialises to

$$[-\partial_L + \gamma(a)]\, G(a, L) = 0 \quad \Longrightarrow \quad G(a, L) = e^{-\gamma(a)L} \qquad \text{(rainbow/ladder RG)} \tag{5.4.24}$$

where the initial condition $G(0, L) = 1$ is imposed by the requirement $\gamma(0) = 0$. The Kilroy case has a nontrivial β-function given by $\beta(a) = -2a\sigma(X(a)) = 2a\gamma(a)$ and the RG equation reads

$$[-\partial_L + 2a\gamma(a)\partial_a - \gamma(a)]\, G(a, L) = 0 \qquad \text{(Kilroy RG)} \tag{5.4.25}$$

5.4.3. Log expansion & renormalisation group (RG) functions. Suppose we are given the solution of a DSE in the combinatorial form

$$X^r(\alpha) = \mathbb{I} + \operatorname{sgn}(s_r)\textstyle\sum_{n\geq 1} c_n^r \alpha^n \tag{5.4.26}$$

with Hopf algebra elements $c_n^r \in \mathcal{H}$ associated with the amplitude $r \in \mathcal{R}$ and coupling α. If we act the renormalised Feynman character on this expression and make use of (5.4.2), we get

$$\begin{aligned}\chi_R(X^r(\alpha)) &= 1 + \operatorname{sgn}(s_r)\textstyle\sum_{n\geq 1}\chi_R(c_n^r)\alpha^n = 1 + \operatorname{sgn}(s_r)\sum_{n\geq 1}\sum_{j\geq 1}\sigma_j^r(c_n)L^j\alpha^n \\ &= 1 + \operatorname{sgn}(s_r)\textstyle\sum_{j\geq 1}\sum_{n\geq 1}\sigma_j^r(c_n)\alpha^n L^j =: 1 + \operatorname{sgn}(s_r)\sum_{j\geq 1}\gamma_j(\alpha)L^j,\end{aligned} \tag{5.4.27}$$

where $\gamma_j^r(\alpha) = \sum_{n\geq j}\sigma_j(c_n^r)\alpha^n$ is a formal power series (remember that $\sigma_j(c_n^r) = 0$ if $n < j$). The resulting expression is the *log expansion* of what we call the *Green's function* of the amplitude $r \in \mathcal{R}$, ie

$$G^r(\alpha, L) = 1 + \operatorname{sgn}(s_r)\textstyle\sum_{j\geq 1}\gamma_j^r(\alpha)L^j =: 1 + \operatorname{sgn}(s_r)\gamma^r(\alpha)\cdot L, \tag{5.4.28}$$

which is the form factor for the amplitude $r \in \mathcal{R}$, where the latter expression is a convenient shorthand that will later come in handy. The formal power series

$$\gamma_j^r(\alpha) = \textstyle\sum_{n\geq j}\sigma_j(c_n^r)\alpha^n \qquad\qquad j = 1, 2, 3, ... \tag{5.4.29}$$

represent what we call the *log-coefficient* or *renormalisation group (RG) functions*. We shall use both terms interchangably in the exposition, depending on which aspect we want to emphasise, ie their being the coefficients of the log expansion or being part of the *renormalisation group (RG) recursion*, to be explicated in the next subsection.

Note that given a Green's function $G^r(\alpha, L)$, these objects are essentially its derivatives with respect to the single-scale momentum parameter $L = \log(-q^2/\mu^2)$, ie

$$\gamma_n^r(\alpha) = (n!)^{-1}\partial_L^n G^r(\alpha, L)|_{L=0}\ . \tag{5.4.30}$$

5.4.4. Renormalisation group (RG) recursion. If we apply $\chi_R = \exp_\star(L\sigma)$ to the combinatorial series (5.4.26), we obtain:

$$\chi_R(X^r(\alpha)) = 1 + \sum_{n\geq 1}\frac{L^n}{n!}\sigma^{\star n}(X^r(\alpha)). \tag{5.4.31}$$

Because this must be equal to the log expansion (5.4.27), we find

$$\gamma_n^r(\alpha) = \frac{1}{n!}\operatorname{sgn}(s_r)\sigma^{\star n}(X^r(\alpha)). \tag{5.4.32}$$

With this result at hand [**Y11**], we are ready for the next

PROPOSITION 5.5 (RG recursion). *All log-coefficient functions $\gamma_n^r(\alpha)$ of the amplitude $r \in \mathcal{R}$ are related to the anomalous dimension $\gamma_1^r(\alpha)$ through the recursion*

$$(n+1)\gamma_{n+1}^r(\alpha) = [\beta(\alpha)\partial_\alpha + \operatorname{sgn}(s_r)\gamma_1^r(\alpha)]\gamma_n^r(\alpha). \tag{5.4.33}$$

PROOF. Analogous to that of Prop.5.4, only that one has to apply $\sigma^{\star n+1} = \sigma \star \sigma^{\star n}$ to coproduct formula (5.4.1):

$$\begin{aligned}\sigma^{\star n+1}(X^r(\alpha)) &= \operatorname{sgn}(s_r)\textstyle\sum_{m\geq 1}[\sigma(X^r(\alpha)) + m\sigma(Q(\alpha))]\sigma^{\star n}(c_m^r)\alpha^m \\ &= [\sigma(X^r(\alpha)) + \sigma(Q)\alpha\partial_\alpha]\sigma^{\star n}(X^r(\alpha)) \\ &= \operatorname{sgn}(s_r)[\operatorname{sgn}(s_r)\gamma_1^r(\alpha) + \beta(\alpha)\partial_\alpha]n!\gamma_n^r(\alpha)\end{aligned} \tag{5.4.34}$$

and the recursion identity follows. □

Concrete examples are the photon RG recursion in QED,

$$(n+1)\gamma^{\sim}_{n+1}(\alpha) = \gamma^{\sim}_1(\alpha)[\alpha\partial_\alpha - 1]\gamma^{\sim}_n(\alpha) \qquad \text{(photon RG recursion)}, \tag{5.4.35}$$

and the Kilroy recursion

$$(n+1)\gamma_{n+1}(a) = \gamma_1(a)[2a\partial_a - 1]\gamma_n(a) \qquad \text{(Kilroy RG recursion)} \tag{5.4.36}$$

while

$$(n+1)\gamma_{n+1}(a) = \gamma_1(a)\gamma_n(a) \qquad \text{(ladder/rainbow RG recursion)} \tag{5.4.37}$$

for the rainbow and ladders has no more information content than (5.4.24) because this recursion implies $\gamma_n(a) = (\gamma_1(a)^n/n!$, where $\gamma_1(a) = -\gamma(a)(= -\gamma_u(a)$ for ladders) is the convention from (5.2.18,5.2.24).

5.5. DSEs in terms of Mellin transforms

We have seen in (5.2.17) how the Mellin transform of the skeleton graph emerges naturally when we try a scaling ansatz for a linear DSE. Although nonlinear DSEs cannot be solved by this method and Mellin transforms consequently do not play the same role, one can recast these equations into a form such that they appear again. The Kilroy approximation (5.2.36) is a nice example which is particularly suited to illustrate how to derive this form of DSE [**Y11, Krei06, KrY06**].

The main trick to be employed is

$$G^r(\alpha, L) = 1 + \operatorname{sgn}(s_r)\textstyle\sum_{n\geq 1}\gamma^r_n(a)L^n = \left[1 + \operatorname{sgn}(s_r)\textstyle\sum_{n\geq 1}\gamma^r_n(a)(-\partial_\rho)^n\right] e^{-\rho L}|_{\rho=0} \tag{5.5.1}$$

which is also true for its inverse, ie $G^r(\alpha, L)^{-1} = G^r(\alpha, -\partial_\rho)^{-1}e^{-\rho L}|_{\rho=0}$ in the case of a propagator amplitude r. This will enable us to recast a DSE into a form that employs the Mellin transforms of the skeletons.

5.5.1. Revisiting Kilroy.

If we take the Kilroy DSE (5.2.40), write $-i\Sigma_K(q) = \not{q}A(q^2)$ and pass over into the Euclidean realm, we obtain

$$A(-q_E^2) = i\frac{a}{2}\int \frac{d^4k_E}{\pi^2}\frac{1}{k_E^2(q_E-k_E)^2}\left[1 - iA(-q_E^2)\right]^{-1}. \tag{5.5.2}$$

The renormalised version of this reads

$$A_R(q_E^2, \mu^2) = i\frac{a}{2}\int \frac{d^4k_E}{\pi^2}\left\{\frac{1}{k_E^2(q_E-k_E)^2} - \frac{1}{k_E^2(\widetilde{q}_E-k_E)^2}\right\}\left[1 - iA_R(q_E^2,\mu^2)\right]^{-1} \tag{5.5.3}$$

with reference momentum $\widetilde{q}_E$ such that $\widetilde{q}_E^2 = \mu^2$. We set $G(a, L) := 1 - iA_R(q_E^2, \mu^2)$ and, using the trick (5.5.1) we arrive at

$$G(a, L) = 1 + \frac{a}{2}\,G(a, -\partial_\rho)^{-1}\int \frac{d^4k_E}{\pi^2}\left\{\frac{1}{k_E^2(q_E-k_E)^2} - \frac{1}{k_E^2(\widetilde{q}_E-k_E)^2}\right\}\left(\frac{k_E^2}{\mu^2}\right)^{-\rho}\Bigg|_{\rho=0}, \tag{5.5.4}$$

in which $L = \ln(q_E^2/\mu^2)$ and $\widetilde{q}_E^2 = \mu^2$ is understood. Now we see that the Mellin transform we had found in (5.2.10) (not miraculously) reappears here:

$$G(a, L) = 1 + a\,G(a, -\partial_\rho)^{-1}(e^{-\rho L} - 1)F(\rho)\big|_{\rho=0} \qquad \text{(Kilroy DSE in Mellin guise)}, \tag{5.5.5}$$

which follows from

$$\int \frac{d^4k_E}{2\pi^2}\frac{1}{k_E^2(q_E-k_E)^2}\left(\frac{k_E^2}{\mu^2}\right)^{-\rho} = \left(\frac{q_E^2}{\mu^2}\right)^{-\rho}F(\rho)\ . \tag{5.5.6}$$

Applying the same method to the rainbow (5.2.14) or the ladder approximation (5.2.21) leads to

(5.5.7)
$$G(a,L) = 1 - a\, G(a,-\partial_\rho)(e^{-\rho L}-1)F_w(\rho)\big|_{\rho=0} \qquad \text{(Rainbow/Ladder DSE in Mellin guise)},$$

where F_w is the Mellin transform of the ladder or the rainbow 1-loop graph of Yukawa theory.

5.5.2. DSE for RG functions. Using the abbreviation $G(a,L) = 1 \pm \gamma(a)\cdot L$, the above DSEs take the form

$$1-\gamma(a)\cdot L = 1 + a\,[1-\gamma(a)\cdot(-\partial_\rho)]^{-1}(e^{-\rho L}-1)F(\rho)\big|_{\rho=0} \qquad \text{(Kilroy DSE)}, \tag{5.5.8}$$

$$1+\gamma(a)\cdot L = 1 - a\,[1+\gamma(a)\cdot(-\partial_\rho)](e^{-\rho L}-1)F_w(\rho)\big|_{\rho=0} \qquad \text{(ladder/rainbow DSE)} \tag{5.5.9}$$

where both make a strong suggestion: that we differentiate them n-times with respect to L and then set $L=0$: doing so, we get

$$n!\gamma_n(a) = -a\,[1-\gamma(a)\cdot(-\partial_\rho)]^{-1}(-\rho)^n F(\rho)\big|_{\rho=0} \qquad \text{(Kilroy)} \tag{5.5.10}$$

$$n!\gamma_n(a) = -a\,[1+\gamma(a)\cdot(-\partial_\rho)](-\rho)^n F_w(\rho)\big|_{\rho=0} \qquad \text{(rainbow/ladder)}, \tag{5.5.11}$$

for the rainbow and the ladder case. These equations, as a system for all $n \geq 1$, constitute an infinite system of coupled formal algebraic equations for the log-coefficient functions.

Let us see what they look like for the anomalous dimension, ie the case $n=1$. Because the Mellin transforms are known in both cases, one can compute the rhs of these equations: for the ladder/rainbow case we find

$$\gamma_1(a) = \frac{a}{c_w}\left[1+\textstyle\sum_{l=1}^{\infty}(-1)^l\; l!\; \gamma_l(a)\right] \qquad \text{(ladder/rainbow)} \tag{5.5.12}$$

where $c_w = 1$ for ladders and $c_w = 2$ for rainbows. This can be easily be combined with the RG recursion (5.4.37) to yield the implicit equation we have seen in Section 5.2:

$$\gamma_1(a) = \frac{a}{2}\left[1+\textstyle\sum_{l=1}^{\infty}(-\gamma_1(a))^l\right] \;\Rightarrow\; 1 = -\frac{a}{c_w\gamma(a)(1-\gamma(a))} = -aF_w(\gamma(a)) \tag{5.5.13}$$

where $\gamma(a) = -\gamma_1(a)$. Surely, this is nothing new. Yet it is worth mentioning that even in complete ignorance of scaling ansatzes, we would have found the solution to the rainbow and ladder approximation by this combination of the RG recursion (5.4.37) and the DSE (5.5.12). However, the Kilroy case leads to a substantially messier expression,

$$\gamma_1(a) = \frac{a}{2}\left[1+\sum_{m\geq 1}\sum_{r\geq m}(-1)^r\; r! \sum_{r_1+\ldots+r_m=r}\gamma_{r_1}(a)\ldots\gamma_{r_m}(a)\right] \qquad \text{(Kilroy)}, \tag{5.5.14}$$

which allows no such simple conclusions. However, perturbatively, one can combine this equation and the RG recursion (5.4.36), to compute the solution recursively to any arbitrary order in the coupling a [**KrY06**].

5.5.3. Photon DSE in terms of Mellin transforms. To see what form this takes in QED for the photon propagator, we apply the same trick at every insertion place that we have performed in (5.5.4) for the rainbow series in Yukawa theory: at the m-th insertion place (where the photon's full propagator sits), we make use of

$$G^{\sim}(\alpha,L)^{-1} = G^{\sim}(\alpha,-\partial_{\rho_m})^{-1}e^{-\rho_m L}\big|_{\rho_m=0} \tag{5.5.15}$$

and obtain

$$
\begin{aligned}
G^{\sim}(\alpha, L) &= 1 - \sum_{j\geq 1} \alpha^j (\chi_R \circ B_+^{q_j^{\sim}})((X^{\sim})^{1-j}) \\
&= 1 - \sum_{j\geq 1} \alpha^j \prod_{l=1}^{j-1} G^{\sim}(\alpha, -\partial_{\rho_l})^{-1} [e^{-(\rho_1+\ldots+\rho_{j-1})L} - 1] F_j(\rho_1, \ldots, \rho_{j-1}) \Big|_{\rho=0} ,
\end{aligned}
\tag{5.5.16}
$$

where $F_j = F_j(\rho_1, ..., \rho_{j-1})$ is the Mellin transform of the Hopf algebra element $q_j^{\sim}$ whose arguments are the regulators of the individual convergence factors that regularise the internal photon lines, ie the Mellin integral contains a regulatorising factor

$$
\prod_{l=1}^{j-1} \left(\frac{k_l^2}{\mu^2}\right)^{-\rho_l}
\tag{5.5.17}
$$

with the effect that there is a regulator for each internal bare photon line l with momentum k_l passing through. Needless to say that the QED case looks like a mess compared to the above examples, altogether approximations.

However, Kreimer and Yeats have combinatorial arguments[12] which one can invoke to prove the existence of a meromorphic function $H_j(\rho)$ of one single argument $\rho \in \mathbb{C}$ such that

$$
\begin{aligned}
\lim_{(\rho_1,\ldots,\rho_{j-1})\to 0} \prod_{l=1}^{j-1} G^{\sim}(\alpha, -\partial_{\rho_l})^{-1} [e^{-(\rho_1+\ldots+\rho_{j-1})L} - 1] F_j(\rho_1, \ldots, \rho_{j-1}) \\
= \lim_{\rho\to 0} G^{\sim}(\alpha, -\partial_\rho)^{1-j} [e^{-\rho L} - 1] H_j(\rho) \ .
\end{aligned}
\tag{5.5.18}
$$

This simplifies (5.5.16):

$$
G^{\sim}(\alpha, L) = 1 - \sum_{j\geq 1} \alpha^j \, G^{\sim}(\alpha, -\partial_\rho)^{1-j} [e^{-\rho L} - 1] H_j(\rho) \big|_{\rho=0}
\tag{5.5.19}
$$

which for the log-coefficient functions takes the form

$$
n!\gamma_n^{\sim}(\alpha) = \sum_{j\geq 1} \alpha^j \, [1 - \gamma^{\sim}(\alpha) \cdot (-\partial_\rho)]^{1-j} (-\rho)^n H_j(\rho) \big|_{\rho=0} \ .
\tag{5.5.20}
$$

If we expand $H_j(\rho) = \sum_{t\geq 0} h_t^{(j)} \rho^{t-1}$ and plug this into (5.5.20), then we get[13] for $n = 1$,

$$
\gamma_1^{\sim}(\alpha) = J(\alpha) + \sum_{m\geq 1} \sum_{r\geq m} r! J_{m,r}(\alpha) \sum_{r_1+\ldots+r_m=r} \gamma_{r_1}^{\sim}(\alpha) \ldots \gamma_{r_m}^{\sim}(\alpha) \qquad \text{(QED)},
\tag{5.5.21}
$$

where the 'skeleton functions' are given by

$$
J(\alpha) := -\sum_{j\geq 1} h_0^{(j)} \alpha^j \ , \qquad J_{m,r}(\alpha) := (-1)^{m+r+1} \sum_{j\geq 1} \binom{1-j}{m} h_r^{(j)} \alpha^j .
\tag{5.5.22}
$$

Notice that (5.5.14) has roughly the same form with

$$
J(a) = \frac{a}{2}, \qquad J_{m,r}(a) = (-1)^r \frac{a}{2} \qquad \text{(Kilroy skeleton functions)}.
\tag{5.5.23}
$$

The highest power of the skeleton functions' expansions signify how many Dyson-Schwinger skeletons one has taken into account.

12 Private communication by Dirk Kreimer.

13 The derivation is relegated to the appendix, Section B.4.

5.6. Nonlinear ordinary differential equations from DSEs

We will in this section see how a DSE in terms of Mellin transforms can be employed to find a nonlinear ordinary differential equation for the corresponding anomalous dimension. The 'inhomogeneity' of this equation turns out to be particularly simple in the case of the Kilroy approximation.

5.6.1. ODE for the Kilroy approximation. The Kilroy DSE in terms of Mellin transforms in combination with the associated RG recursion allows us to formulate an ordinary differential equation (ODE) which can be transformed into an implicit transcendental equation, as the next assertion informs us [**KrY06**].

PROPOSITION 5.6 (Kilroy ODE). *The RG recursion and the DSE of the Kilroy approximation imply the nonlinear ODE*

$$\gamma(a) + \gamma(a)(2a\partial_a - 1)\gamma(a) = \frac{a}{2} \tag{5.6.1}$$

for the anomalous dimension $\gamma(a)$.

PROOF. The Kilroy DSE (5.5.10) and (5.2.10) entail

$$\gamma_1(a) + 2\gamma_2(a) = a\,[1 - \gamma(a)\cdot(-\partial_\rho)]^{-1}\rho(1-\rho)F(\rho)\big|_{\rho=0} = \frac{a}{2} \tag{5.6.2}$$

while on the other hand, the RG recursion (5.4.36) implies

$$\gamma_1(a) + 2\gamma_2(a) = \gamma_1(a) + \gamma_1(2a\partial_a - 1)\gamma_1(a), \tag{5.6.3}$$

where $\gamma(a) = \gamma_1(a)$ is the anomalous dimension. □

Modern computer algebra software[14] can nowadays transform the ODE (5.6.1) into the implicit equation

$$\sqrt{\frac{a}{\pi}}e^{-Z(a)} = 1 + \operatorname{erf}(Z(a)), \tag{5.6.4}$$

which we have seen in Section 5.2. As already mentioned there, this result and the ODE (5.6.1) have been published by Broadhurst and Kreimer in [**BroK01**], but with different conventions[15].

QED case. The general case, of course, does not do us the favour of yielding such a nice and simple differential equation. However, any linear combination of the first two log-coefficient functions yield a first order differential equation. Including higher log-coefficient functions leads to higher order ODEs, ie equations with higher-order derivatives. For QED, it turns out that the choice

$$P(\alpha) := \gamma_1^{\sim}(\alpha) - 2\gamma_2^{\sim}(\alpha) \tag{5.6.5}$$

is particularly interesting as it leads to the ODE [**BKUY09, Y11**]

$$\gamma_1^{\sim}(\alpha) + \gamma_1^{\sim}(\alpha)(1 - \alpha\partial_\alpha)\gamma_1^{\sim}(\alpha) = P(\alpha). \tag{5.6.6}$$

which at first order perturbation theory in $P(\alpha)$ is explicitly solvable and has a *genuinely* non-pertubative solution! This is going to be the theme of the next chapter.

Before we close this chapter, we take (5.5.20) and use it to write $P(\alpha)$ is terms of Mellin transforms

$$P(\alpha) = \gamma_1^{\sim}(\alpha) - 2\gamma_1^{\sim}(\alpha) = \sum_{j\geq 1}\alpha^j\,[1 - \gamma^{\sim}(\alpha)\cdot(-\partial_\rho)]^{1-j}\rho(1+\rho)H_j(\rho)\big|_{\rho=0}\ , \tag{5.6.7}$$

which shows that it extracts data from the photon's quenched skeleton expansion.

---

[14] We have used Maple 16.

[15] One has to do the replacements $a \to a/2$ and $\gamma \to -\gamma/2$ to see that both results agree.

CHAPTER 6

Landau pole and flat contributions in quantum electrodynamics

This chapter is a pedagogical account and review of the results published previously in [**KlaKrei13**] which build upon the work of Kreimer and Yeats who first introduced (5.6.6) and then studied it in [**BKUY09**] and a variant of it in [**BKUY10**] pertaining to QCD. Amongst other aspects discussed there, the focus was to draw conclusions about the beta function and a possible Landau pole from the asymptotic behaviour of the function $P(\alpha)$.

We augment these investigations in Section 6.1 by discussing how possible zeros of this function may be decisive for the Landau pole question. Section 6.2 introduces the concept of *flat contributions* which refers to flat functions, ie functions with a vanishing Taylor series.

Through this property, these contributions are a nonperturbative feature. Yet the concept of 'flat contributions' turns out to be mathematically rather vague which is why we shall carefully discern the situations when it can be made precise. However, when the rhs of the 'photon equation' (5.6.6) has a flat contribution, then, not surprisingly on account of the RG recursion (5.4.35), the anomalous dimension must also have one. Interestingly, on account of the nonlinearity of (5.6.6), the converse is not true and, furthermore, perturbing the rhs flatly can only result in a flat perturbation of the solution.

The subsequent sections investigate a toy model with $P(\alpha) = \alpha$ which may be seen as a perturbative approximation of the photon equation (5.6.6) as long as the coupling is small. The resulting equation can be solved analytically and its solution, presented in Section 6.3, exhibits a nonperturbative flat contribution, making it all the more worth studying.

We show in Section 6.4 that this flat contribution of the toy model hampers the growth of the beta function but not to the extent that a Landau pole is averted. However, in the light of this aspect, the possibility of avoiding a Landau pole through such nonperturbative contributions seems plausible even in proper QED. Section 6.5 scrutinises the effect of the flat part on the location of the Landau pole.

We present the resulting toy model photon self-energy in Section 6.6 and study the impact the flat contribution exerts on the Green's function. Although we *can* relate the perturbative series of the function $P(\alpha)$ to the skeleton diagram expansion (5.6.7), we *cannot* find a canonical diagrammatic interpretation of our toy photon self-energy in terms of a *resummation scheme* like in the case of renormalon chain [**FaSi97, Ben99**] or rainbow approximations [**DeKaTh97, DelEM97, KrY06**]. Our approach is of a fundamentally different nature: we take a nonperturbative equation, solve it perturbatively for the first loop order and yet get an instantonic (that is, flat) contribution.

6.1. Photon equation and Landau pole criterion

6.1.1. Landau pole. Quantum field theory teaches us that the coupling parameter of a renormalisable theory is *scale-dependent.* This refers either to the length or to the momentum scale. We shall focus on the latter and write the *renormalsation group equation* for *running coupling* $\alpha(L)$ in QED as

$$\frac{\partial}{\partial L}\alpha(L) = \beta(\alpha(L)), \qquad \text{(RG equation)} \tag{6.1.1}$$

We take it as being dependent on the momentum variable $L = \ln(-q^2/\mu^2)$ with Minkowski momentum $q \in \mathbb{M}$ and renormalsation reference parameter μ^2. The function $\beta(x)$ is the beta function, as introduced in Proposition 5.4.

DEFINITION 6.1. *We say that QED has a Landau pole at a point $L_* \in \mathbb{R}$ if $\alpha(L_*) = \infty$.*

Whether or not such a pole exists depends on the behaviour of the beta function. What we can tell from experiments, the beta function increases monotonically and $\beta(x) > 0$. If we integrate (6.1.1), we find

$$\int_{\alpha_0}^{\alpha} \frac{dx}{\beta(x)} = L - L_0, \tag{6.1.2}$$

in which $\alpha_0 = \alpha(L_0)$ is some (experimentally determined) reference coupling. Suppose the beta function grows so quickly that

$$\int_{\alpha_0}^{\infty} \frac{dx}{|\beta(x)|} < \infty, \tag{6.1.3}$$

ie that the integral converges for the limit $\alpha \to \infty$. Then $L_* = L_0 + \int_{\alpha_0}^{\infty} \frac{dx}{\beta(x)}$ would be the Landau pole.

6.1.2. Existence of a Landau pole. We shall from now on write the photon equation (5.6.6) in the form

$$\gamma(\alpha) + \gamma(\alpha)(1 - \alpha\partial_\alpha)\gamma(\alpha) = P(\alpha), \tag{6.1.4}$$

with $\gamma(\alpha) := \gamma_1^{\sim}(\alpha)$ being the anomalous dimension of the photon. We will see that it habours a sufficient criterion for the existence of a Landau pole. We take a slightly different view from that in [**BKUY09**] where

$$\mathcal{L}(P) = \int_{x_0}^{\infty} \frac{2dx}{x(\sqrt{1+4P(x)} - 1)} < \infty \qquad (x_0 > 0) \tag{6.1.5}$$

is found to be a necessary and sufficient criterion for the existence of a Landau pole. Unfortunately, we do not know how $P(\alpha)$ behaves. But what do we know about $P(\alpha)$?

First note that from perturbation theory we know $\beta(\alpha) > 0$, and hence $P(\alpha) > 0$ for small $\alpha > 0$ by how it is defined. This function may have zeros: at a point $\alpha_0 \in (0, \infty)$, where $P(\alpha_0) = 0$ we see that by

$$0 = P(\alpha_0) = \gamma(\alpha_0)[1 + \gamma(\alpha_0) - \alpha_0\gamma'(\alpha_0)] \tag{6.1.6}$$

we have either $\gamma(\alpha_0) = 0$ and thus $\beta(\alpha_0) = \alpha_0\gamma(\alpha_0) = 0$ or

$$1 + \gamma(\alpha_0) - \alpha_0\gamma'(\alpha_0) = 0. \tag{6.1.7}$$

We exclude the possibility that $P(\alpha)$ has an infinite number of zeros and compare the following two assumptions from a physical point of view:

(H1) *$P(\alpha)$ has no nontrivial zero, ie no zero other than $\alpha_0 = 0$.*
(H2) *The anomalous dimension $\gamma(\alpha)$ has no nontrivial zero whereas $P(\alpha)$ does have a finite number of zeros.*

Notice that (H1) implies $\gamma(\alpha) > 0$ for all $\alpha > 0$: for small $\alpha > 0$, both factors in (6.1.6) are positive and since $P(\alpha)$ never vanishes for $\alpha > 0$, none of these factors can ever change sign. The next two propositions will help us decide which of the two assumptions is stronger.

PROPOSITION 6.2 (Asymptotics of anomalous dimension I). *Assume (H1), ie that $P(\alpha)$ vanishes nowhere except at the origin. Then there exist a constant $A > 0$ and a point $\alpha_* \in \mathbb{R}_+$ such that*

$$\gamma(\alpha) < A\alpha - 1 \qquad \forall \alpha > \alpha_*, \tag{6.1.8}$$

which entails that the beta function is dominated on (α_*, ∞) by a quadratic polynomial in α.

Proof. Pick any $\alpha_* \in (0, \infty)$ and set $A := [1 + \gamma(\alpha_*)]/\alpha_*$. By (6.1.6), assumption (H1) implies

$$\forall \alpha \in \mathbb{R}_+ : \ \gamma'(\alpha) < \frac{1+\gamma(\alpha)}{\alpha} \quad \Rightarrow \quad \gamma'(\alpha_*) < \frac{1+\gamma(\alpha_*)}{\alpha_*} = A \tag{6.1.9}$$

and thus by definition, the linear function $h_A(\alpha) := A\alpha - 1$ is the line that meets $\gamma(\alpha)$ at the point $\alpha = \alpha_*$ but has stronger growth there. Hence $h(\alpha) := \gamma(\alpha) - h_A(\alpha)$ satisfies $h(\alpha_*) = 0$ and $h'(\alpha_*) < 0$. Consequently, there is an $\varepsilon > 0$ such that $h(\alpha) < 0$ for all $\alpha \in (\alpha_*, \alpha_* + \varepsilon)$. For a sign change of $h(\alpha)$ there must be a point $\overline{\alpha} > \alpha_*$ where $h(\overline{\alpha}) < 0$ and $h'(\overline{\alpha}) = 0$ which implies the contradiction

$$0 = h'(\overline{\alpha}) = \gamma'(\overline{\alpha}) - A < \frac{1+\gamma(\overline{\alpha})}{\overline{\alpha}} - A = \frac{1+\gamma(\overline{\alpha}) - A\overline{\alpha}}{\overline{\alpha}} = \frac{h(\overline{\alpha})}{\overline{\alpha}}. \tag{6.1.10}$$

□

Note that the asymptotics of (6.1.8) does not touch on the Landau pole question of QED: the growth of the beta function may or may not be strong enough for a Landau pole to exist. Regarding the second assumption (H2), we will see that we need the extra property that $P(\alpha) > 0$ for large enough $\alpha \in \mathbb{R}_+$, ie behind the last zero on the half-line $\mathbb{R}_+$, not to have a Landau pole enforced.

Proposition 6.3 (Asymptotics of anomalous dimension II). *Suppose (H1), ie $\gamma(\alpha)$ vanishes nowhere other than at the origin and $P(\alpha)$ has a finite number of zeros. Let furthermore $P(\alpha)$ be such that it becomes negative for sufficiently large α, ie there is an $\alpha_* \in \mathbb{R}_+$ with $P(\alpha) < 0$ for all $\alpha \geq \alpha_*$. Then there exists a constant $A > 0$ such that*

$$\gamma(\alpha) > A\alpha - 1 \tag{6.1.11}$$

for all $\alpha > \alpha_*$ and QED has a Landau pole because the anomalous dimension $\gamma(\alpha)$ grows too rapidly for large α.

Proof. With A defined as above, the assumptions imply

$$\gamma'(\alpha_*) > \frac{1+\gamma(\alpha_*)}{\alpha_*} = A > 0 \tag{6.1.12}$$

and hence this time $h_A(\alpha) = A\alpha - 1$ is the line that meets $\gamma(\alpha)$ at the point $\alpha = \alpha_*$ but is growing less there. Consequently, $h(\alpha) = \gamma(\alpha) - h_A(\alpha)$ satisfies $h(\alpha_*) = 0$ and $h'(\alpha_*) > 0$. Now, there is an $\varepsilon > 0$ such that $h(\alpha) > 0$ for all $\alpha \in (\alpha_*, \alpha_* + \varepsilon)$. For a sign change of $h(\alpha)$ there must be a point $\overline{\alpha} > \alpha_*$ where $h(\overline{\alpha}) > 0$ and $h'(\overline{\alpha}) = 0$ which implies the contradiction

$$0 = h'(\overline{\alpha}) = \gamma(\overline{\alpha}) - A > \frac{1+\gamma(\overline{\alpha})}{\overline{\alpha}} - A = \frac{h(\overline{\alpha})}{\overline{\alpha}}. \tag{6.1.13}$$

The second assertion holds because of

$$\left| \int_{\alpha_*}^{\infty} \frac{dx}{\beta(x)} \right| \leq \left| \int_{\alpha_*}^{\infty} \frac{dx}{x h_A(x)} \right| < \infty, \tag{6.1.14}$$

where one should note that $0 < h_A(x) = Ax - 1 \leq \gamma(x)$ for all $x \geq \alpha_*$. □

In summary, QED can only be free of a Landau pole if $P(\alpha) \geq 0$ for large enough α and

- if we assume (H1), that $P(\alpha)$ has no nontrivial zeros, then there may or may not be a Landau pole,
- but assuming (H2), ie that $P(\alpha)$ has nontrivial zeros, we have to make the additional assumption that $P(\alpha)$ never becomes negative, ie that the number of zeros is even to come to the same conclusion.

Assumption (H1) is therefore in some sense weaker than (H2). With the latter proposition and the definition of $P(\alpha)$, we arrive at

COROLLARY 6.4 (Landau pole criterion). *QED has a Landau pole if $\gamma_1(\alpha) < 2\gamma_2(\alpha)$ for large enough $\alpha > 0$. A necessary condition for the non-existence of a Landau pole is therefore given by*

$$\gamma_1(\alpha) \geq 2\gamma_2(\alpha) \qquad \text{for sufficiently large } \alpha, \tag{6.1.15}$$

ie the second log-coefficient function must not win out over the first.

This means a necessary condition for the avoidance of a Landau pole is that the contributions to the first momentum-log power must dominate those to the second power for large couplings.

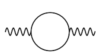
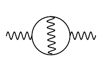
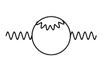
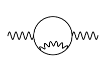
6.1.3. Perturbative Expansion. To study the photon equation (6.1.4) perturbatively with respect to $P(\alpha)$, we expand this function in α. Let us see what the first coefficients are. The RG functions' perturbation series are

$$\gamma_j(\alpha) = -\sigma_j(\mathbb{I} - \sum_{k\geq 1} c_k\alpha^k) = \sum_{k\geq 1}\sigma_j(c_k)\alpha^k = \sum_{k\geq j}\sigma_j(c_k)\alpha^k =: \sum_{k\geq j}\gamma_k^{(j)}\alpha^k, \tag{6.1.16}$$

where $\gamma_k^{(j)} = \sigma_j(c_k)$ is the k-th (asymptotic) Taylor coefficient of $\gamma_j(\alpha)$. Note that $c_k \in \mathcal{H}$ is a linear combination of k-loop vacuum polarization graphs and cannot contribute to log-powers greater than k, hence $\sigma_j(c_k) = 0$ for $j > k$. This entails that the asymptotic expansion of $\gamma_j(\alpha)$ starts with the j-th coefficient which is also implied by the photon's RG recursion in (5.4.35). However, with these maps at hand, we draw on (5.6.5) and see that the perturbative expansion of the function $P(\alpha)$ is given by

$$P(\alpha) = \sum_{k\geq 1}[\sigma_1(c_k) - 2\sigma_2(c_k)]\alpha^k. \tag{6.1.17}$$

Note that $\chi_R(c_1) = \sigma_1(c_1)L$ and $\chi_R(c_2) = \sigma_1(c_2)L + \sigma_2(c_2)L^2$. From the results

(6.1.18) $\chi_R(c_1) = \chi_R($  $) = \frac{L}{3\pi},\quad \chi_R(c_2) = \chi_R($  $+$  $+$  $) = \frac{L^2}{4\pi^2},$

found in [**GoKaLaSu91**], we read off the coefficients of $P(\alpha)$ to find

$$P(\alpha) = \frac{1}{3}\frac{\alpha}{\pi} - \frac{1}{2}\left(\frac{\alpha}{\pi}\right)^2 + \mathcal{O}(\alpha^3). \tag{6.1.19}$$

Given the perturbative series of $P(\alpha)$, the ODE in (6.1.4) determines the anomalous dimension $\gamma(\alpha)$ perturbatively: let $u_1, u_2, ...$ be the asymptotic coefficients of $\gamma(\alpha)$ and $r_1, r_2, ...$ those of $P(\alpha)$, then (6.1.4) imposes

$$r_k = u_k + \sum_{l=2}^{k-1}(1-l)u_l u_{k-l} \tag{6.1.20}$$

giving a nice recursion [**Y13**],

$$u_1 = r_1\ ,\ \ u_2 = r_2\ ,\ \ u_3 = r_3 + r_2 r_1\ ,\ \ u_4 = r_4 + 2r_2 r_1^2 + 2r_1 r_3 + r_2^2,\ \dots,\text{etc.} \tag{6.1.21}$$

Though we know next to nothing about $P(\alpha)$, we are confident that its Taylor series is divergent, yet probably still Gevry-1, that is, its Borel transform

$$\mathscr{B}P(\alpha) = \sum_{k\geq 1}\frac{r_k}{k!}\alpha^k \tag{6.1.22}$$

should have nonvanishing radius of convergence[1], which entails that it has a holomorphic continuation $\mathtt{cont}_I(\mathscr{B}P)(\alpha)$ to some small interval $I \subset \mathbb{R}_+$. But we do not know whether $P(\alpha)$ is Borel-summable in the sense that $I = \mathbb{R}_+$ and that the Borel-Laplace transform

$$\mathscr{L}[\mathscr{B}P](\alpha) = \int_0^\infty dt\, e^{-t}\mathtt{cont}(\mathscr{B}P)(\alpha t), \tag{6.1.23}$$

the so-called *Borel sum* of $P(\alpha)$, gives us back $P(\alpha)$.

We tend to believe this is not the case. The reason is that we expect so-called *flat functions* to play a role. These functions most likely preclude (6.1.23) from delivering back $P(\alpha)$ because they do not satisfy the conditions of Watson's theorem (see Appendix Section A.8). Here is a definition of flat functions.

Definition 6.5 (Flat functions). *A smooth function $f : (0, \infty) \to \mathbb{R}$ is called* flat *if it has a vanishing Taylor series at zero.*

The prime example well known to physicists is $f(\alpha) = \exp(-1/\alpha)$. It is usually referred to as an *instanton contribution*. Although QED exhibits no instantons, $P(\alpha)$ may nevertheless have a *transseries representation* which does indeed feature such functions. We shall not elaborate on this point and defer a discussion of transseries to the next chapter.

Clearly, the recursion (6.1.20) is blind to such contributions. What this means is that any flat function added to $P(\alpha)$ lies in the same germ of analytic functions. Perhaps not surprisingly, [**BKUY10**] have found an upper bound for the difference of two different solutions of (6.1.4) in terms of a flat function:

Proposition 6.6. *Let $P \in \mathscr{C}^2(\mathbb{R}_+)$ be positive with $P(0^+) = 0$ and $P'(0^+) \neq 0$. Then two solutions γ and $\widetilde{\gamma}$ of the ODE (6.1.4) differ by a flat function, more precisely,*

$$|\gamma(\alpha) - \widetilde{\gamma}(\alpha)| \leq E\alpha \exp(-F/\alpha), \qquad \forall \alpha \in [0, \alpha_0] \tag{6.1.24}$$

where the the constants $E, F > 0$ depend on $\alpha_0 > 0$.

Proof. See Theorem 5.1 in [**BKUY10**] . □

To round off this section, we mention for the sake of completeness that a more general version of (6.1.4) pertaining not just to QED has been studied in [**BKUY09**] with the following result about the function $P(\alpha)$. On the assumption that

(A1) $P(\alpha)$ is twice differentiable and
(A2) strictly positive on $(0, \infty)$ with $P(0) = 0$,

then a global solution exists iff $J(P) := \int_{x_0}^\infty dz P(z) z^{-3}$ converges for some $x_0 > 0$. If furthermore

(A3) $P(\alpha)$ is everywhere increasing,

then there exists a *'separatrix'*: a global solution that separates all global solutions from those existing only on a finite interval. We shall see in Section 6.3 how this latter situation arises in the 1-loop approximation with respect to $P(\alpha)$ and ensures that the family of solutions covers the whole set of solutions.

Moreover, it turns out that in this toy model, the separatrix picks out the very physical solution that corresponds to a beta function whose growth is weakest among those of all other possible physical cases. Although it is not weak enough to avert a Landau pole, it may very well be in the case of the 'true' $P(\alpha)$.

[1] Readers unfamiliar with Borel summation should consult Appendix Section A.8.

6.2. Flat contributions

For the sake of a neater notation, we set $D := 1 - \alpha\partial_\alpha$ and

$$\mathscr{F} := \left\{ f \in \mathscr{C}^\infty(\mathbb{R}_+) \mid \forall n \in \mathbb{N} : \lim_{\alpha \downarrow 0} \partial_\alpha^n f(\alpha) = 0 \right\} \tag{6.2.1}$$

be the algebra of all flat functions. To study flat contributions, we have to make an assumption about $P(\alpha)$ against our better judgement due to a mathematical subtlety. We still believe our results to be of value. The issue is this: because the algebra $\mathscr{F}$ is a subspace in the space of smooth functions $\mathscr{C}^\infty(\mathbb{R}_+)$, there exists a projector $\pi_\mathscr{F} : \mathscr{C}^\infty(\mathbb{R}_+) \to \mathscr{F}$ such that $f \in \mathscr{C}^\infty(\mathbb{R}_+)$ can be decomposed into

$$f = (\mathrm{id} - \pi_\mathscr{F})(f) + \pi_\mathscr{F}(f) = f_0 + f_1, \tag{6.2.2}$$

where $f_1 := \pi_\mathscr{F}(f) \in \mathscr{F}$ is flat. However, there is surely not just one projector and hence not just one possible decomposition of f into 'flat' and 'non-flat': take any flat $g \in \mathscr{F}$, then

$$f = f_0 - g + f_1 + g = \widetilde{f_0} + \widetilde{f_1} \tag{6.2.3}$$

with $\widetilde{f_0} := f_0 - g$ being the non-flat and $\widetilde{f_1} = f_1 + g$ being the flat part. As a consequence, there is no unique decomposition of the desired kind and things get cloudy at the attempt to find a strict mathematical definition of 'flat contribution'. However, this is not so if we restrict ourselves to the subspace of functions $f \in \mathscr{C}^\infty(\mathbb{R}_+)$ with convergent Taylor series

$$(Tf)(\alpha) := \sum_{k\geq 0} \frac{1}{k!} f^{(k)}(0^+)\alpha^k \tag{6.2.4}$$

at zero which have an analytic continuation $\mathtt{cont}(Tf)$ to the half-line $[0,\infty)$. An easy example is

$$f(\alpha) = \frac{1}{1+\alpha} + e^{-1/\alpha}. \tag{6.2.5}$$

Its Taylor series $\sum_{j\geq 0}(-1)^j\alpha^j$ is convergent, enjoys an analytic continuation to $[0,\infty)$ and yet it converges nowhere to $f(\alpha)$. We denote the algebra of these functions by $\mathscr{M}$ and define the projector $\pi_\mathscr{F} : \mathscr{M} \to \mathscr{F}$ as the (uniquely determined) linear operator that subtracts the analytic continuation of the convergent Taylor series from the function, ie

$$\pi_\mathscr{F}(f) := f - \mathtt{cont}(Tf) = (\mathrm{id} - \mathtt{cont} \circ T)f \tag{6.2.6}$$

is an element in $\mathscr{F}$ and the decomposition $f = \mathtt{cont}(Tf) + \pi_\mathscr{F}(f)$ is unique. We therefore have the decomposition

$$\mathscr{M} = \mathscr{M}_0 \oplus \mathscr{F} \tag{6.2.7}$$

with $\mathscr{M}_0 := (\mathrm{id} - \pi_\mathscr{F})\mathscr{M}$ being also an algebra. We write $\pi_{\mathscr{M}_0} := (\mathrm{id} - \pi_\mathscr{F}) = \mathtt{cont} \circ T$ for its projector. In fact, $\mathscr{M}_0$ is the well known algebra of analytic functions on $[0,\infty)$. It is invariant under differential operators and hence a *differential algebra*. In particular, this implies D-invariance, ie $D\mathscr{M}_0 \subset \mathscr{M}_0$. The flat algebra also has this property. For later reference, we list its properties:

(i) $\mathscr{F}$ is D-invariant, that is, $D\mathscr{F} \subseteq \mathscr{F}$.
(ii) The product of any function in $\mathscr{M}$ and a flat function is flat: $\mathscr{M}\mathscr{F} \subset \mathscr{F}$, ie $\mathscr{F}$ is an ideal in the algebra $\mathscr{M}$.

In summary, $\mathscr{M}$ is the class of functions $f \in \mathscr{C}^\infty(\mathbb{R}_+)$ with convergent Taylor series (at $\alpha = 0$) *that do not converge to* f only if $\pi_\mathscr{F}(f) \neq 0$, ie if f features a nontrivial flat part (which renders it non-analytic). Note that the operator D has the one-dimensional kernel $\ker D = \mathbb{R}\alpha \subset \mathscr{M}_0$ and we therefore have a third property:

(iii) If $f \in \mathbb{R}\alpha + \mathscr{F}$, then $Df \in \mathscr{F}$.

We shall draw on (i)-(iii) in the proofs of the following assertions which we view as interesting on the following grounds.

Being aware that $\gamma(\alpha)$ and almost surely also $P(\alpha)$ are non-analytic functions with divergent Taylor series, we would like to point out that within our approach of approximating $P(\alpha)$ perturbatively by a polynomial in α and hence by a function in the class $\mathscr{M}_0$, it makes perfect sense to us to assume that $\gamma(\alpha)$ is at most in the class $\mathscr{M}$: the coefficient recursion in (6.1.20) can then only be expected to lead to a divergent series of $\gamma(\alpha)$ in mathematically contrived situations. By allowing $P(\alpha)$ to be in $\mathscr{M}$, we go one doable step *beyond* perturbation theory. 'Doable' because the decomposition (6.2.7) is mathematically well-defined in a way that it otherwise would not be: it enables us to get a grip on the otherwise vague concept of flat contributions which the β-function may or may not feature.

Claim 6.7 (Flat perturbations I). *Let $P \in \mathscr{M}$ with a nontrivial flat part: $\pi_{\mathscr{F}}(P) \neq 0$. Then any solution of the ODE*

$$\gamma(\alpha) + \gamma(\alpha) D\gamma(\alpha) = P(\alpha) \tag{6.2.8}$$

does also have a nontrivial flat part, that is, $\pi_{\mathscr{F}}(\gamma) \neq 0$.

Proof. Let $\pi_{\mathscr{F}}(\gamma) = 0$. Then follows $\pi_{\mathscr{F}}(D\gamma) = 0$ and also $\pi_{\mathscr{F}}(\gamma D\gamma) = 0$ by $\mathscr{M}_0$ being a differential algebra. This entails $\pi_{\mathscr{F}}(P) = 0$. ☐

Note that the converse is not true: even in cases where $\pi_{\mathscr{F}}(P) = 0$ may a flat part pop up in the anomalous dimension γ. The reason is that in such cases, (6.2.8) implies

$$\pi_{\mathscr{F}}(\gamma) + \pi_{\mathscr{F}}(\gamma D\gamma) = 0 \tag{6.2.9}$$

which can be massaged into a differential equation for the flat function $\pi_{\mathscr{F}}(\gamma)$ and has solutions beyond the trivial one. The next assertion treats the toy case $P(\alpha) \in \alpha + \mathscr{F}$ which is equivalent to $\pi_{\mathscr{F}}(P) = 0$ and reveals how the anomalous dimension γ is affected.

Claim 6.8 (Flat perturbations II). *Let $\gamma(\alpha)$ be a solution of*

$$\gamma(\alpha) + \gamma(\alpha) D\gamma(\alpha) = \alpha. \tag{6.2.10}$$

and $\overline{\gamma}(\alpha)$ the solution of its flatly perturbed version

$$\overline{\gamma}(\alpha) + \overline{\gamma}(\alpha) D\overline{\gamma}(\alpha) = \alpha + f(\alpha), \tag{6.2.11}$$

where $f \in \mathscr{F}$. Then $\overline{\gamma} - \gamma \in \mathscr{F}$, ie a flat perturbation of the rhs of (6.2.10) leads to a flat perturbation of its solution.

Proof. We draw on the result of the next section: according to (6.3.2), any solution of (6.2.10) satisfies

$$\gamma(\alpha) \in \alpha + \mathscr{F} \tag{6.2.12}$$

and thus $\pi_{\mathscr{M}_0}(\gamma) = \alpha$ and we write $\gamma = \alpha + \pi_{\mathscr{F}}\gamma$. Because $P(\alpha)$ determines the perturbation series of $\gamma(\alpha)$ uniquely through (6.1.20) where flat parts do not participate, the perturbation series of $\overline{\gamma}$ and γ coincide. Trivially, this means that if γ has a convergent Taylor series, so does $\overline{\gamma}$. Hence $\gamma, \overline{\gamma} \in \mathscr{M}$ and there is a decomposition $\overline{\gamma} - \gamma = h_0 + h_1$, where $h_0 := \pi_{\mathscr{M}_0}(\overline{\gamma} - \gamma)$ and $h_1 \in \mathscr{F}$ is flat. We will show that the function h_0 vanishes everywhere. Subtracting (6.2.10) from (6.2.11) yields

$$\begin{aligned} &\overline{\gamma} - \gamma + (\overline{\gamma} - \gamma) D\overline{\gamma} + \gamma D(\overline{\gamma} - \gamma) \\ &\qquad = (h_0 + h_1) + (h_0 + h_1) D(\alpha + \pi_{\mathscr{F}}(\gamma) + h_0 + h_1) + (\alpha + \pi_{\mathscr{F}}\gamma) D(h_0 + h_1) = f. \end{aligned} \tag{6.2.13}$$

To get rid of all the flat stuff, we apply the projector $\pi_{\mathscr{M}_0} = (\mathrm{id} - \pi_{\mathscr{F}})$ to both sides of this equation and obtain

$$h_0 + h_0 D h_0 + \alpha D h_0 = 0, \tag{6.2.14}$$

where we have used $D\alpha = 0$ and drawn on (i)-(iii). We can rewrite (6.2.14) in the form

$$(h_0 + \alpha) + (h_0 + \alpha)D(h_0 + \alpha) = \alpha, \tag{6.2.15}$$

and find that $\gamma = h_0 + \alpha$. This means $h_0 \in \mathscr{F}$ and therefore $\pi_{\mathscr{F}}(h_0) = h_0$. Consequently, on account of $\pi_{\mathscr{F}}(h_0) = 0$ by $h_0 \in \mathscr{M}_0$, we see that $h_0 = 0$. □

The next proposition generalises this latter assertion.

PROPOSITION 6.9 (Flat Perturbations III). *Let $P \in \mathscr{M}_0$ be such that $P(0^+) = 0$, $P'(0^+) \neq 0$ and $P(\alpha) > 0$ for $\alpha > 0$. Then any solution $\overline{\gamma}$ of*

$$\overline{\gamma}(\alpha) + \overline{\gamma}(\alpha)D\overline{\gamma}(\alpha) = P(\alpha) + f(\alpha) \tag{6.2.16}$$

with $f \in \mathscr{F}$ differs from a solution of $\gamma(\alpha) + \gamma(\alpha)D\gamma(\alpha) = P(\alpha)$ only flatly, ie $\overline{\gamma} - \gamma \in \mathscr{F}$.

PROOF. Let again $\overline{\gamma} - \gamma = h_0 + h_1$ be the decomposition as in Claim 6.8 and $\gamma_0 = \pi_{\mathscr{M}_0}(\gamma)$ the non-flat part of γ. Then follows $\pi_{\mathscr{M}_0}(\overline{\gamma}) = \gamma_0 + h_0$. Purging both ODEs of all flat contributions by applying $\pi_{\mathscr{M}_0}$ yields

$$\gamma_0 + \gamma_0 D\gamma_0 = P \qquad \text{and} \qquad (\gamma_0 + h_0) + (\gamma_0 + h_0)D(\gamma_0 + h_0) = P, \tag{6.2.17}$$

where $\pi_{\mathscr{M}_0}(P) = P$ by assumption. Then we have $h_0 \in \mathscr{F}$ by Proposition 6.6, ie the result from [**BKUY10**]. Since $h_0 \in \mathscr{M}_0$ by definition, we conclude $h_0 = 0$. □

6.3. First order non-analytic approximation

We draw on (6.1.18), set $c := 1/(3\pi)$ and choose $P(\alpha) = c\alpha$ for a first order approximation. Now, the photon equation (6.1.4) takes the form

$$\gamma(\alpha) + \gamma(\alpha)(1 - \alpha\partial_\alpha)\gamma(\alpha) = c\alpha. \tag{6.3.1}$$

This equation has already been studied in [**BKUY09**] where the reader can find a plot of the direction field for the anomalous dimension $\gamma(\alpha)$. We shall review their results and expound them in somewhat more detail. It is an easy exercise to prove that

$$\gamma(\alpha) = c\alpha[1 + W(\xi e^{-\frac{1}{c\alpha}})] = c\alpha + c\alpha W(\xi e^{-\frac{1}{c\alpha}}) \in \mathbb{R}\alpha + \mathscr{F} \tag{6.3.2}$$

provides a family of solutions in which $W(x)$ is the famous *Lambert W function*, defined as the inverse function of $x \mapsto x\exp(x)$, and the family parameter $\xi := (\gamma(1/c) - 1)e^{\gamma(1/c)}$ is fixed by the initial condition for $\gamma(\alpha)$ at $\alpha = 1/c$. This follows from

$$\gamma(1/c) - 1 = W(\xi e^{-1}) = \xi e^{-1}e^{-W(\xi e^{-1})} = \xi e^{-[1+W(\xi e^{-1})]} = \xi e^{-\gamma(1/c)}, \tag{6.3.3}$$

in turn a consequence of

$$x = W(x)e^{W(x)}. \tag{6.3.4}$$

This function has two branches, denoted by W_0 and W_{-1} (see Figure 1) and emerges in physics whenever the identity (6.3.4) may be exploited to solve a transcendental equation[2].

We shall ignore the second branch $W_{-1}(x)$, for the following reason: it is only defined on the half-open interval $[-1/e, 0)$ and coerces us to choose $\xi < 0$. On this interval, it rapidly drops down an abyss where one finds $W_{-1}(0^-) = -\infty$. But although it turns out that $\gamma(0^+) = 0$ in this branch, one finds $\gamma(\alpha) < 0$ for all couplings which entails $\beta(\alpha) = \alpha\gamma(\alpha) < 0$ for the beta function. As this is not what we would call QED-like behaviour, we discard this branch. In contrast to this, we will see that the first branch $W_0(x)$ serves our purposes perfectly well. We will denote it by $W(x)$.

[2]See for example the QCD-related papers [**GaKaG98**] and [**Nest03**].

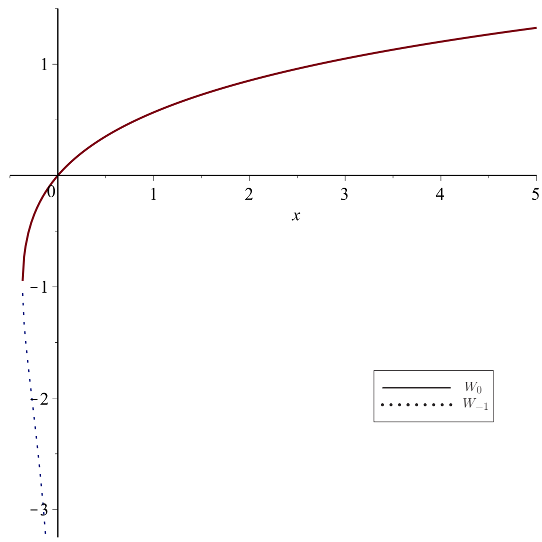


FIGURE 1. The two branches of the Lambert W function. Note that the second branch W_{-1} (dotted line) is restricted to the interval $[-1/e, 0)$ and vanishes nowhere.

Because our approximation does only hold for very small values of the coupling parameter α and, for example

$$\frac{c}{137} W(\xi e^{-\frac{137}{c}}) \sim 10^{-562} \tag{6.3.5}$$

with $\xi = -e^{-1}$, that is, the flat part is practically invisible. We shall scale away c such that $c\alpha \to \alpha$ without renaming of functions and view all of the following results arising from $P(\alpha) = \alpha$ as those of an interesting toy model.

6.3.1. Beta function. The point $\xi_* := -e^{-1}$ turns out to be critical for the beta function

$$\beta(\alpha) = \alpha\gamma(\alpha) = \alpha^2[1 + W(\xi e^{-\frac{1}{\alpha}})]. \tag{6.3.6}$$

The only way the beta function can vanish at some point $\alpha_0 \in (0, \infty)$ is when

$$W(\xi e^{-\frac{1}{\alpha_0}}) = -1 \tag{6.3.7}$$

which by $x = W(x)e^{W(x)}|_{W=-1} = -1e^{-1}$ implies

$$\xi e^{-\frac{1}{\alpha_0}} = -e^{-1} = \xi_* \ . \tag{6.3.8}$$

This means $\xi < \xi e^{-\frac{1}{\alpha_0}} = \xi_*$ and that if we choose $\xi < \xi_*$, then the zero is at

$$\alpha_0 = \frac{1}{1 + \ln|\xi|}. \tag{6.3.9}$$

The limit $\xi \uparrow \xi_*$ throws this point to infinity. The initial condition which corresponds to the choice $\xi = \xi_*$ is given by

$$\gamma(1) = 1 + W(-e^{-2}) \approx 0.841 \tag{6.3.10}$$

and characterizes the *separatrix* beta function $\beta^*(\alpha) = \alpha\gamma_1^*(\alpha)$. The choice $\xi < \xi_*$ entails a nontrivial zero $\alpha_0 > 0$ but is somewhat *unphysical*: their solution $\beta(\alpha)$ simply ceases to exist at α_0 and has a divergent derivative at this point, ie $\beta(\alpha_0) = 0$ and $\beta'(\alpha_0) = -\infty$. We therefore conclude that only $\xi \geq \xi_* = -e^{-1}$ is physically permissible for further consideration. Note that the usual one-loop approximation for the beta function corresponds to the case $\xi = 0$, which is also physical. Figure 2 shows examples for different choices of ξ.

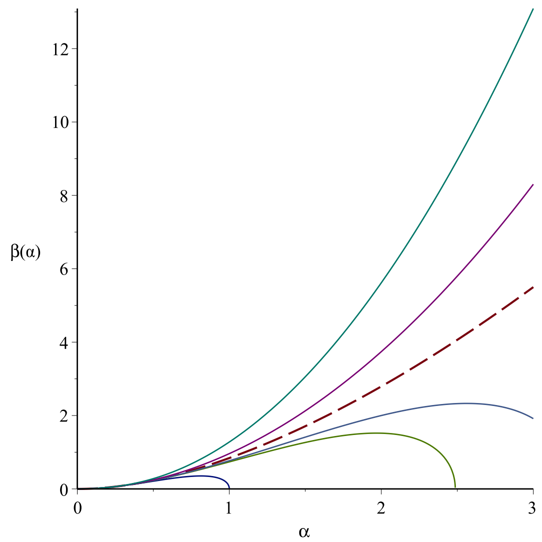


FIGURE 2. The beta function for different choices of ξ: only the separatrix corresponding to $\xi = \xi_*$ (dashed line) and the curves above it with $\xi > \xi_*$ are physical, whereas those with a zero for $\xi < \xi_*$ are not.

The possible solutions reflect the results of [**BKUY09**] alluded to in the previous section: the separatrix $\beta^*(\alpha)$ separates global solutions from those with a finite interval of definition. As a result, the above family of solutions in (6.3.2) covers the set of all possible solutions.

6.4. Landau pole avoidance

If we insert $P(x) = x$ into the integral of (6.1.5), we see that $\mathcal{L}(P) < \infty$ is satisfied which, according to [**BKUY09**], means that the toy model has a Landau pole. We can in fact see this more explicitly, because one can simply integrate the RG equation (6.1.1) of the running coupling to give

$$L - L_0 = \int_{\alpha_0}^{\alpha} \frac{dx}{\beta(x)} = \ln \left| \frac{W(\xi e^{-1/\alpha})}{W(\xi e^{-1/\alpha_0})} \right| \tag{6.4.1}$$

with reference coupling $\alpha(L_0) = \alpha_0$. We find that our model has a Landau pole at $L = L_*$ since the integral

$$L_* - L_0 = \int_{\alpha_0}^{\infty} \frac{dx}{\beta(x)} = \ln \left| \frac{W(\xi)}{W(\xi e^{-1/\alpha_0})} \right| = \frac{1}{\alpha_0} + W(\xi e^{-1/\alpha_0}) - W(\xi) \tag{6.4.2}$$

exists for any $\alpha_0 > 0$ and $\alpha(L)$ diverges for a finite $L = L_*$. To avoid a Landau pole, we require that this very integral diverge which in our case means that the beta function must not grow too rapidly. For the separatrix choice $\xi = \xi_* = -e^{-1}$ we find by expanding the Lambert W function

$$\beta(\alpha) \sim \sqrt{2}\alpha^{\frac{3}{2}} - \frac{2}{3}\alpha + \mathcal{O}(1) \qquad \text{as} \qquad \alpha \to \infty \tag{6.4.3}$$

and thus a decreased growth compared to the instanton-free 1-loop beta function given by $\beta(\alpha)|_{\xi=0} = \alpha^2$ which is because the *instantonic contribution works towards the avoidance of a Landau pole* by means of the asymptotics

$$1 + W(-e^{-1-\frac{1}{\alpha}}) \sim \sqrt{\frac{2}{\alpha} - \frac{2}{3\alpha}} + \mathcal{O}(\alpha^{-2/3}) \qquad \text{as} \qquad \alpha \to \infty. \tag{6.4.4}$$

This is an example in which the instantonic contribution alters the convergence behaviour of the integral

$$\int_{x_0}^{\infty} \frac{dx}{\beta(x)} \tag{6.4.5}$$

and may thus in other cases exclude the existence of a Landau pole, notwithstanding that any perturbative series of the beta function is blind to such contributions.

Given the above facts about the photon equation and the prominent role of the flat algebra $\mathscr{F}$, it is not unreasonable to assume that the anomalous dimension $\gamma(\alpha)$ and hence the beta function

$$\beta(\alpha) = \alpha\gamma(\alpha) = \beta_0(\alpha) + \beta_1(\alpha) \tag{6.4.6}$$

has a flat piece $\pi_{\mathscr{F}}(\beta) = \beta_1$. However, let us now be really bold and assume that this part takes the form

$$\beta_1(\alpha) = (\bar{\beta}(\alpha) - \beta_0(\alpha))e^{-\frac{r}{\alpha}}, \tag{6.4.7}$$

where $r > 0$ is some positive real number and $\bar{\beta}(\alpha)$ is such that $\int_R^\infty \bar{\beta}(x)^{-1}dx = \infty$ for any $R > 0$. Let us furthermore assume that the non-flat piece $\beta_0 = \pi_{\mathscr{A}_0}(\beta)$ satisfies $\lim_{t\downarrow 0}(1-e^{-rt})\beta_0(1/t) = 0$. Then, on account of

$$\beta(\alpha) = \beta_0(\alpha) + (\bar{\beta}(\alpha) - \beta_0(\alpha))e^{-\frac{r}{\alpha}} \sim \bar{\beta}(\alpha) \qquad \text{as} \qquad \alpha \to \infty \tag{6.4.8}$$

one has

$$\int_{\alpha_0}^{\infty} \frac{dx}{\beta(\alpha)} = \infty \tag{6.4.9}$$

and thus a Landau pole-free theory. However, these conditions seem a bit contrived and, alas, we do not know the real beta function of QED and it is an inherent feature of perturbation theory with respect to the coupling that this integral *converges*[3]. Therefore, any approximation of the running coupling as a solution of the RG equation (6.1.1) within this framework is bound to have a pole which, however, we hardly need to remind the reader, leaves the question of a Landau pole of the real theory untouched.

6.5. Landau pole of the toy model

We carry on with our toy model and solve (6.4.1) for the running coupling $\alpha(L)$ to find

$$\alpha_\xi(L) = \frac{\alpha_0}{1 - \alpha_0(L - L_0) + g_\xi(\alpha_0, e^{L-L_0})}, \tag{6.5.1}$$

where $g_\xi(\alpha_0, x) := \alpha_0(1-x)W(\xi e^{-1/\alpha_0})$ is flat in its first argument $\alpha(L_0) = \alpha_0$. We shall now have a look at the running coupling for both spacelike and timelike photons and compare the instanton-free case $\xi = 0$ with the separatrix case $\xi = \xi_*(= -e^{-1})$. The hampering effect of the flat contribution on the growth of the beta function turns out to result in lower values of the coupling in the case of large momentum transfer. This is to be expected as the beta function determines the momentum scale dependence of the coupling through the RG equation $\partial_L\alpha(L) = \beta(\alpha(L))$.

6.5.1. Spacelike photons. Figure 3 has a plot of the running coupling $\alpha_\xi(L)$ with $\xi = \xi_*$ and the instanton-free case $\xi = 0$ for spacelike photons where $L \in \mathbb{R}$ due to $-q^2 > 0$ and reference coupling $\alpha_0 = 0.1$. The diagram shows that the flat piece $g_\xi(\alpha_0, e^L)$ shifts the Landau pole from $L' = 1/\alpha_0$ to $L'' = L' + ...$, the solution of the transcendental equation

$$L'' = L' + (1 - e^{L''})W(-e^{-1-1/\alpha_0}). \tag{6.5.2}$$

[3]In the case of zeros of the beta function choose α_0 beyond them.

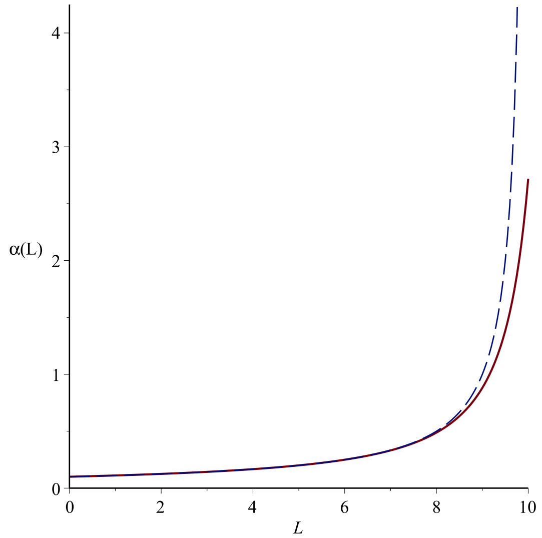


FIGURE 3. Spacelike photons: the running coupling $\alpha_\xi(L)$ for $\xi = \xi_*$ (solid) and the instanton-free case $\xi = 0$ (dashes) at reference coupling $\alpha_0 = 0.1$ for $L_0 = 0$.

Note that $L'' > L'$ due to $(1 - e^L)W(-e^{-1-1/\alpha_0}) > 0$ for any $L > 0$ and that on account of the transcendentality of (6.5.2) there is no canonical way to define a reference scale, usually denoted by Λ.

6.5.2. Timelike photons. In the case of timelike photons, when $-q^2 < 0$, the running coupling in (6.5.1) has an imaginary part, we write (6.5.1) in the form

$$\widetilde{\alpha}_\xi(-s) := \alpha_\xi(\log(-s)) = \frac{\alpha_0}{1 - \alpha_0 \log(-s) + g_\xi(\alpha_0, -s)}, \tag{6.5.3}$$

where $L = \log(-s)$ with $s := q^2/\mu^2$ and $L_0 = 0$ for reference point $s_0 = -1$. Complex-valued 'timelike' couplings have been studied in QCD: [**PenRo81**] have argued that $|\widetilde{\alpha}(-s)|$ is to be favoured over $\widetilde{\alpha}(|-s|)$ as perturbation coupling parameter for timelike processes. They find better agreement with experimental results at lower order of perturbation theory.

Whether or not this pertains to QED, we have plotted this parameter in Figure 4 for the two cases $\xi = \xi_*$ and $\xi = 0$.

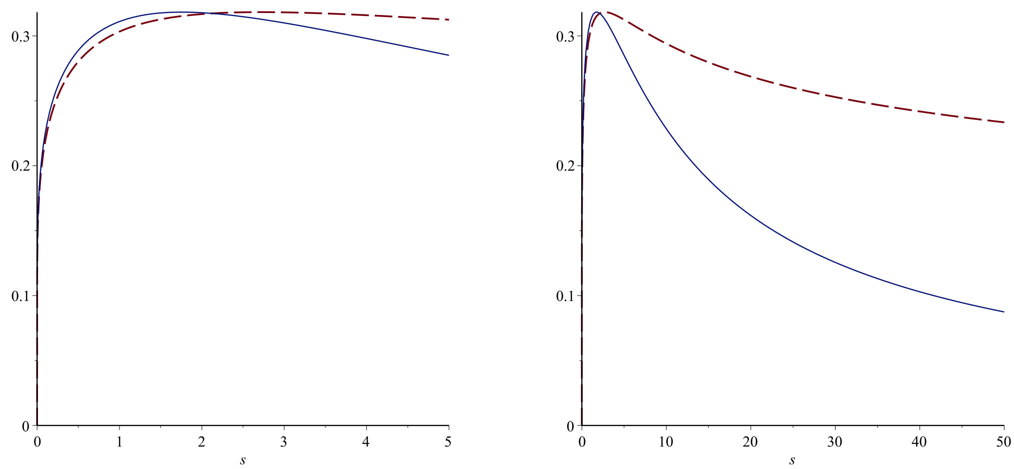


FIGURE 4. Coupling parameter $|\widetilde{\alpha}_\xi(-s)|$ for timelike photons at lower (left) and higher (right) momenta: $\xi = \xi_*$ (solid) and $\xi = 0$ (dashes) with reference coupling $\alpha_0 = 1$.

Because the coupling obviously sports a branch cut, we consider the dispersion relation

$$\widetilde{\alpha}_\xi(-s) = \int_0^\infty \frac{\Omega_\xi(\eta)}{s-\eta} d\eta \tag{6.5.4}$$

and calculate the spectral density by taking the limit

$$\lim_{\varepsilon\downarrow 0} \{\widetilde{\alpha}_\xi(-x-i\varepsilon) - \widetilde{\alpha}_\xi(-x+i\varepsilon)\} = -2\pi i \Omega_\xi(x). \tag{6.5.5}$$

yielding

$$\Omega_\xi(\eta) = \frac{\alpha_0^2}{[1-\alpha_0\log(\eta) + g_\xi(\alpha_0,\eta)]^2 + (\alpha_0\pi)^2}, \tag{6.5.6}$$

which equals the square modulus of the timelike coupling in (6.5.3): $\Omega_\xi(\eta) = |\widetilde{\alpha}_\xi(-\eta)|^2$. Apart from the absolute value, the spectral density has a graph of the same shape as that in Figure 4, where we see that the instantonic contribution shows a significant effect at strong reference coupling $\alpha_0 = 1$: beyond the pair-creation 'bump', higher mass state contributions are suppressed.

The seemingly natural interpretation of this for timelike photons in terms of a weaker interaction in the s-channel where electrons and positrons annihilate has to be more than taken with a pinch of salt: though these deviations seem to be significant, we have to remind ourselves that these are toy model results. As pointed out in Section 3, we cannot expect our (single flavour) toy QED to hold for large couplings around $\alpha_0 = 1$.

To make the impact of the flat contribution visible, however, we have to go up to this level of the coupling strength and accept that we at the same time enter the realm of a toy model: the implicit assumption of choosing this reference coupling is that the running coupling is still given by (6.5.3).

Both in the case of timelike and spacelike photons, it is by no means far-fetched to conclude that if flat contributions impede the beta function's growth, the running coupling may exhibit lower values at higher momentum transfer also in a real-world (3-flavour) QED.

6.6. Photon self-energy

Let us recall that the renormalised photon self-energy

$$\Pi(\alpha, e^L) := \gamma(\alpha)\cdot L = \sum_{j\geq 1} \gamma_j(\alpha) L^j \tag{6.6.1}$$

in terms of its log-expansion has its correct place inside the transversal part of the full renormalised photon propagator

$$\Pi_{\mu\nu}(q) = \frac{g_{\mu\nu} - q_\mu q_\nu/q^2}{q^2[1-\Pi(\alpha, -q^2/\mu^2)]}\ , \tag{6.6.2}$$

in massless QED with Minkowski metric $g_{\mu\nu}$ in mostly minus signature and renormalisation point μ^2 in momentum subtraction scheme.

If we take the anomalous dimension in (6.3.2) setting $c = 1$, apply the RG recursion (5.4.35) and calculate all higher log-coefficients, we find for the self-energy an interesting result which we present in the next

Claim 6.10. *Given the exact solution $\gamma(\alpha)$ of the nonlinear toy model ODE (6.3.1), the recursion (5.4.35) yields*

$$j!\gamma_j(\alpha) = \alpha W(\xi e^{-\frac{1}{\alpha}}) \qquad j\geq 2, \tag{6.6.3}$$

that is, only flat higher log-coefficient functions and

$$\Pi_\xi(\alpha, e^L) = \alpha L + \alpha(e^L - 1)W(\xi e^{-\frac{1}{\alpha}}) = \alpha\ln(-q^2/\mu^2) - \alpha\left(1+q^2/\mu^2\right) W(\xi e^{-\frac{1}{\alpha}}) \tag{6.6.4}$$

for the photon self-energy.

Proof. We proceed by induction. First $j = 2$.

$$2!\gamma_2(\alpha) = \gamma(\alpha)(\alpha\partial_\alpha - 1)\gamma(\alpha) \overset{(6.3.1)}{=} \gamma(\alpha) - \alpha = \alpha W(\xi e^{-\frac{1}{\alpha}}). \tag{6.6.5}$$

Note that this is flat. Let now $j \geq 2$. Then

$$\begin{aligned} (j+1)!\gamma_{j+1} &\overset{(5.4.35)}{=} \gamma(\alpha\partial_\alpha - 1)j!\gamma_j = \gamma(\alpha\partial_\alpha - 1)\alpha W = \gamma(\alpha\partial_\alpha - 1)(\gamma - \alpha) \\ &= \gamma(\alpha\partial_\alpha - 1)\gamma = 2\gamma_2 = \alpha W, \end{aligned} \tag{6.6.6}$$

where we have used $(\alpha\partial_\alpha - 1)\alpha = 0$ in the fourth step. For the self-energy then follows

$$\Pi_\xi(\alpha, e^L) = \gamma \cdot L = \gamma L + \sum_{j\geq 2} \gamma_j L^j = \alpha(1+W)L + \sum_{j\geq 2} \frac{1}{j!}\alpha W L^j = \alpha L + \alpha(e^L - 1)W \tag{6.6.7}$$

and thus the result. ☐

In the notation of the previous section we set $s = q^2/\mu^2$ with Minkowski momentum $q \in \mathbb{M}$ and write

$$\Pi_\xi(\alpha, -s) = \alpha \log(-s) - g_\xi(\alpha, -s) = \alpha \log(-s) - (1+s)\alpha W(\xi e^{-1/\alpha}). \tag{6.6.8}$$

To see how the instantonic contribution affects the renormalised propagator, we define

$$P_\xi(\alpha, -s) := \frac{1}{s[1 - \Pi_\xi(\alpha, -s)]} = \frac{1}{s[1 - \alpha\log(-s) + g_\xi(\alpha, -s)]} \tag{6.6.9}$$

and study its properties for $\xi = \xi_*$ and in particular how the flat contribution causes this quantity to deviate from its instanton-free version $\Pi_0(\alpha, -s) = \Pi_\xi(\alpha, -s)|_{\xi=0}$.

6.6.1. Spacelike photons. For spacelike photons, the Green's function is real-valued due to $S := -s = -q^2/\mu^2 > 0$ and leads to the propagator

$$P_\xi(\alpha, S) = -\frac{1}{S[1 - \alpha\log(S) + g_\xi(\alpha, S)]}. \tag{6.6.10}$$

To see the flat contribution's impact, we compare this quantity with $P_0(\alpha, S) = P_\xi(\alpha, S)|_{\xi=0}$. Figure 5 has plots of the squares $|P_0(\alpha, S)|^2$ and $|P_\xi(\alpha, S)|^2$ displaying two aspects:

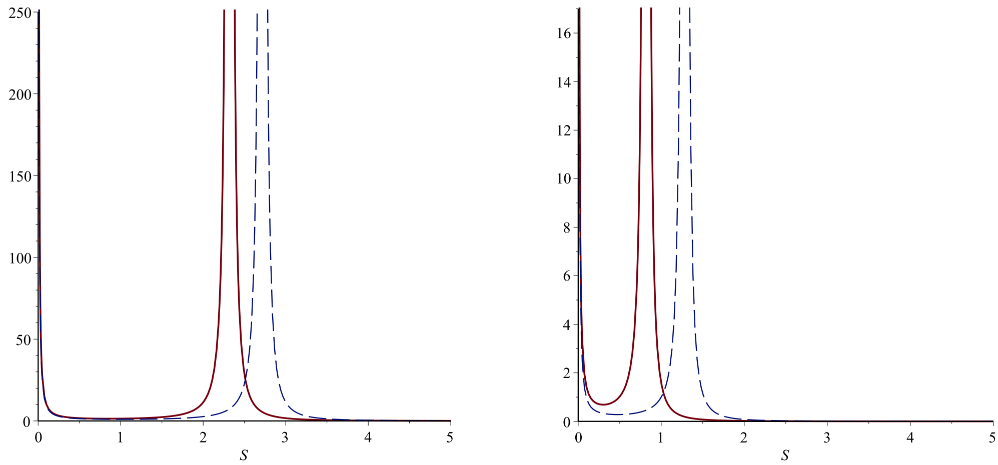


Figure 5. Pole shift of the propagator squares $|P_0(\alpha, S)|^2$ (dashed line), $|P_\xi(\alpha, S)|^2$ (solid line) for $\xi = \xi_*$ and spacelike photons at $\alpha = 1$ (left) and $\alpha = 4$ (right).

- firstly, in both cases ($\xi = 0$ and $\xi = -1/e = \xi_*$), the propagator exhibits a pole which is situated at higher momenta in the weak coupling regime (left diagram of Figure 5) than in the strong coupling regime (right diagram of Figure 5)

- secondly, the instantonic contribution causes a *pole shift* towards lower momenta, where this effect is more pronounced at larger and negligible at lower values of the coupling.

However, since these poles are those of a toy model, we need not interprete them.

6.6.2. Timelike photons: Källén-Lehmann spectral function. For timelike photons, where $-q^2 < 0$ and thus $-s < 0$, the propagator has a branch cut on the real axis. The spectral function $\rho_\xi(\alpha, \omega)$ in the Källén-Lehmann spectral form of the propagator is given by

$$P_\xi(\alpha, -s) = \frac{1}{s[1 - \Pi_\xi(\alpha, -s)]} = \frac{1}{s} + \int_0^\infty d\omega \, \left(\frac{1}{s - \omega} + \frac{1}{1 + \omega} \right) \rho_\xi(\alpha, \omega), \tag{6.6.11}$$

where the integrand has been chosen so as to warrant the renormalisation condition

$$s\Pi(\alpha, -s)|_{s=-1} = 1. \tag{6.6.12}$$

To extract the spectral function, we compute the limit

$$\lim_{\varepsilon \downarrow 0} \{P_\xi(\alpha, -x - i\varepsilon) - P_\xi(\alpha, -x + i\varepsilon)\} = -2\pi i \rho_\xi(\alpha, x) \tag{6.6.13}$$

for $x > 0$ and obtain the Källén-Lehmann spectral function

$$\rho_\xi(\alpha, \omega) = \frac{\alpha}{\omega} \frac{1}{[1 - \alpha \ln \omega + g_\xi(\alpha, -\omega)]^2 + (\alpha\pi)^2}. \tag{6.6.14}$$

Figure 6 shows a plot of the spectral function $\rho(\alpha, \omega) := \rho_{\xi_*}(\alpha, \omega)$ for the separatrix solution at different coupling strengths α. Notice that the dispersion integral has no trouble converging.

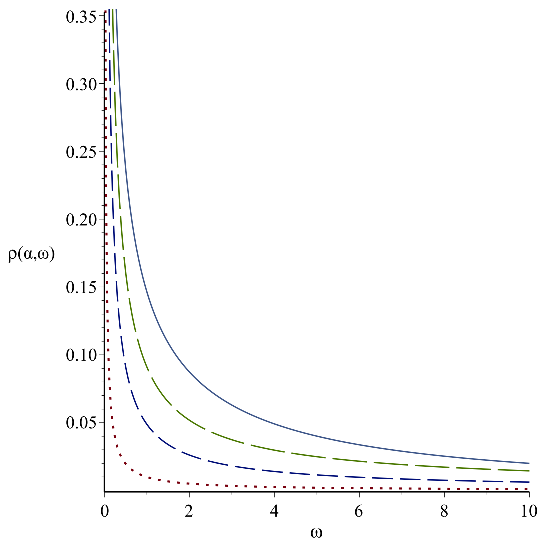


FIGURE 6. Spectral function for various coupling strengths: $\alpha = 0.01$ (dots), $\alpha = 0.05$ (dashes), $\alpha = 0.1$ (long dashes) and $\alpha = 0.5$ (solid).

For large $\omega \to \infty$, the integrand decreases rather fast:

$$\begin{aligned} \xi \neq 0: \quad & \left(\frac{1}{s - \omega} + \frac{1}{1 + \omega} \right) \rho_\xi(\alpha, \omega) \sim -\frac{\omega^{-5}}{\alpha W(\xi e^{-1/\alpha})^2} \\ \xi = 0: \quad & \left(\frac{1}{s - \omega} + \frac{1}{1 + \omega} \right) \rho_0(\alpha, \omega) \sim \frac{\omega^{-4}}{\alpha (\ln \omega)^2}, \end{aligned} \tag{6.6.15}$$

where the effect of the flat contribution is nicely visible, no matter how flat and invisible it may be! For the lower limit of the integration, $\omega \downarrow 0$, we have

$$\left(\frac{1}{s-\omega}+\frac{1}{1+\omega}\right)\rho_\xi(\alpha,\omega) \sim \frac{1+s}{s}\frac{1}{\omega(\ln\omega)^2} \tag{6.6.16}$$

for all $\xi \geq \xi_* = -1/e$. On account of the primitive

$$\int \frac{d\omega}{\omega(\ln\omega)^2} = -\frac{1}{\ln\omega}, \tag{6.6.17}$$

we find that the integrand of the dispersion integral in (6.6.11) is also well-behaved at the lower integration bound.

However, apart from the fact that we are dealing with a toy model here, we have considered *massless* QED, and can thus not expect our spectral function to encapsulate any valid physics below the pair-creation threshold $\omega_0 \approx 4m^2$.

6.6.3. Instantonic contribution. Apart from changing the asymptotics of the spectral function for large masses, the alterations brought about by the nonperturbative contribution are less pronounced at lower masses $\omega > 0$, yet still visible at stronger couplings. To see it, we compare the spectral function $\rho_\xi(\alpha,\omega)$ with its instanton-free version $\rho_0(\alpha,\omega)$ to see this effect. As the diagrams of Figure 7 show for $\alpha = 5$, the flat contribution leads to a slightly increased contribution of lower mass states. For large masses ω there is only a small change towards a smaller contribution.

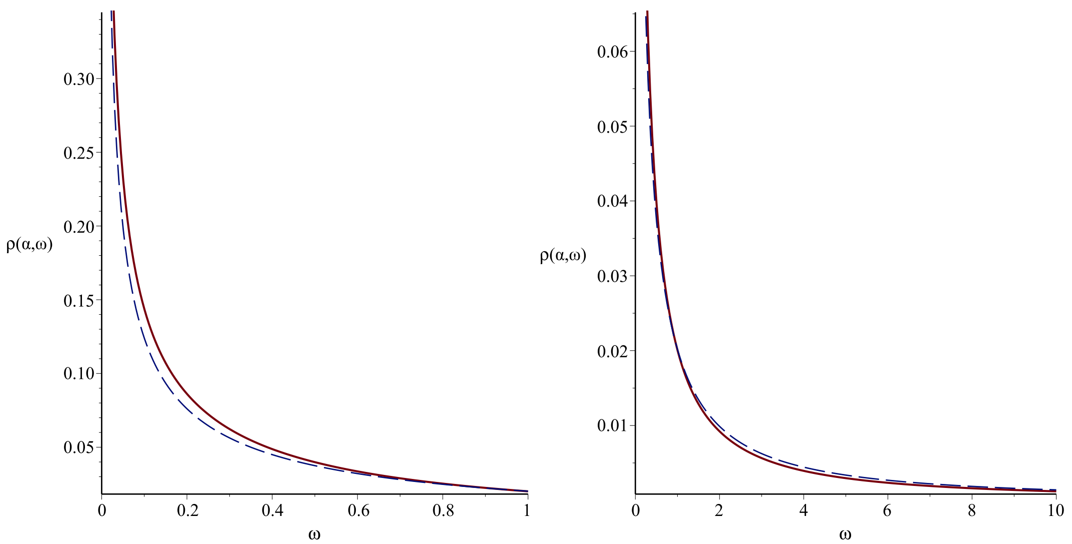


FIGURE 7. The two spectral functions $\rho_0(\alpha,\omega)$ (dashed line), $\rho(\alpha,\omega)$ (solid line) for $\alpha = 5$ and small/large mass contributions (diagram on the left/right).

However, for large masses ω the function $g_\xi(\alpha,-\omega)$ dominates over the logarithmic part in the denominator of $\rho(\alpha,\omega)$ in (6.6.14) and suppresses higher mass contributions much more than the logarithmic contribution by itself, as can be seen nicely in the asymptotics shown in (6.6.15). Interestingly enough, due to the fact that $g_\xi(\alpha,-\omega)$ will dominate over any polynomial in the variable $L = \ln\omega$ for large ω, this picture would not change qualitatively if we took higher loop contributions into account.

CHAPTER 7

Resurgent transseries and Dyson-Schwinger equations

It is widely known that the physics of renormalisable quantum field theories is not entirely captured by perturbation theory. A clear indication of this is given by the general asymptotic behaviour of the coefficients, leading almost surely to a non-convergent, that is, *asymptotic series*. Consequently, perturbation theory cannot be sufficient to define the observables of a QFT and one cannot do without nonperturbative methods.

A question of especial import is whether perturbation theory may still habour some nonperturbative information. This issue has, for example, been addressed by Dunne and Ünsal for the energy levels of two quantum-mechanical systems, the double well and the periodic Sine-Gordon model [**DunUen14**]. The good news is, they were able to answer it in the affirmative, even to the extent that the coefficients of the perturbative series contain all nonperturbative information necessary to determine these functions.

Looking at their interesting work, one has to add: *perturbation theory alone has no nonperturbative tales to tell!* To make a connection to the nonperturbative world, they impose a *boundary condition*, an equation of necessarily nonperturbative character. Needed therefore, in particular in QFT, are *nonperturbative conditions* in the form of nonperturbative equations like the ones we have studied in the foregoing chapters.

That is not say that perturbative coefficients have no nonperturbative information, but when a perturbative series formally satisfies such nonperturbative conditions without us knowing anything of them, we are simply not able to extract the nonperturbative information enclosed in its coefficients. But even if a nonperturbative equation is known, plugging a perturbative series in as an ansatz leads to no nonperturbative insight.

However, the situation changes drastically if one uses so-called *resurgent transseries*, in some sense a generalisation of power series. And herein lies the trick: these series have perturbative and nonperturbative parts whose coefficients will in general be subjected to conditions once such ansatz is inserted into a nonperturbative equation.

This strategy has recently been employed in the context of string theory [**CoSaESVo15**], but to our knowledge not yet to the nonperturbative equations we discussed extensively in previous chapters, namely Dyson-Schwinger equations (DSEs) and their associated renormalisation group (RG) recursions. This is what we shall do in this last chapter: we investigate these equations using a transseries ansatz.

The objective is to check whether these equations impose sufficient conditions on the anomalous dimension and the higher RG functions such that the coefficients of their conjectured resurgent transseries disclose nonperturbative information. In fact, we shall see that under certain provisions to be explicated in this chapter, this is indeed the case. Moreover, we can make an equally strong statement as Dunne and Ünsal about the systems under scrutiny in this chapter and conclude that for them THE PERTURBATIVE SECTOR DETERMINES THE NONPERTURBATIVE SECTOR COMPLETELY!

Section 7.1 introduces the concept of transseries as generalised formal expansions and explains the canonical algebraic structures on the set of such series, including a derivation that we shall need later. However, we can only give a rough idea as to why and how these series represent the RG functions, the central observables in this work. The reason is that to this date no ones knows whether transseries really are capable of capturing the physics of these quantities and what their exact form is. But physicists currently have high hopes of resurgent transseries and believe the evidence is clearly in favour of this optimistic view [**Sti02, AnSchi15**]

Anyhow, this does not deter us in the slightest to adopt the working assumption that they belong to the class of so-called *resurgent functions* which, according the *theory of resurgent functions* as devised by Écalle [**Eca81**], really do enjoy representations as transseries expansions.

Starting from this assumption, we take the liberty of simply treating transseries as algebraic objects to toy around with, whilst still keeping in the back of our mind the idea that *if* the observables of a renormalisable QFT actually do fall into this class of functions, then this is more than a mathematical game.

In Section 7.2 we take the view that the RG recursion gives rise to a *discrete dynamical system*, its discrete-time flow being steered through a specific subset of transseries. The transseries representing the RG functions will in this picture inhabit the discrete orbit.

Because we aim at watching the flow of perturbative and nonperturbative information as it is driven by the dynamical system, we introduce in Section 7.3 transseries with abstract coefficients, where we have chosen an algebraic language. Along the way, we devise some useful and in our mind straightforward terminology to keep track of the flow of perturbative and nonperturbative data.

Although somewhat idiosycratic, we deem it nevertheless a good and apt terminology on the grounds that it enables us to see in Section 7.4 that the discrete RG flow preserves one specific feature of the transseries in its orbit and thereby warrants a certain orderliness in how perturbative and nonperturbative data is being passed on.

The algebraic formulation will finally pay off in Section 7.5 where we present the main assertions regarding the paradigms of this work: the rainbow and ladder, as well as the Kilroy approximation. And, finally, the anomalous dimension of the photon in QED.

In all these cases, the coefficients of the transseries are required to satisfy nonlinear difference equations which encode the principle of 'THE PERTURBATIVE DETERMINES THE NONPERTURBATIVE'.

7.1. Resurgent transseries for quantum field theory

The general definition of *transseries* is rather technical and requires some formal mathematical machinery that we shall blithely avoid here for a simple reason: we do not need it in this work, since we are only dealing with so-called *resurgent transseries*, that is, a very special case. Although we will give in the following a vague idea of the general form of transseries, we refer the reader to the mathematical literature [**Ed09, Hoe06**] for a thorough treatment.

7.1.1. Transseries and transmonomials. Let us start by considering some examples to get the idea. The formal expansions

$$\begin{aligned} \frac{\log z}{(1-z)(\log z - e^z)} &= \sum_{n,m\geq 0} (e^z)^m z^n (\log z)^{-m} \\ \left(1+e^{z+e^{-z}}\right)^{-1} &= \sum_{n\geq 0} (-1)^n (e^z)^n (e^{e^{-z}})^n \end{aligned} \tag{7.1.1}$$

and

$$1 + \frac{1}{2^z} + \frac{1}{3^z} + \frac{1}{4^z} + ... = \sum_{n\geq 1} (e^{-z})^{\ln n} \tag{7.1.2}$$

are all examples of transseries. The first two belong to the important class of *grid-based transseries*[1], whereas the latter, which of course represents the famous Riemann zeta function, does not. The various powers of the exponentials, logarithms and monomials in z and their products in these expansions, viewed as pure symbols, are what one calls *transmonomials.* Loosely speaking, transseries are then formal series of the form

$$f = \sum_{\mathfrak{g} \in \mathfrak{T}} f_{\mathfrak{g}} \mathfrak{g} \tag{7.1.3}$$

where the sum extends over all transmonomials $\mathfrak{g}$ of a given set of transmonomials $\mathfrak{T}$, and the coefficients are (for the most part) real-valued: $f_{\mathfrak{g}} \in \mathbb{R}$.

Things become utterly fancy when towered exponentials of transseries like (7.1.3) to any tower order (!) are also contained in the set $\mathfrak{T}$, which, at least in part, is what the theory of transseries is about [**Ed09, Hoe06**]. The reader can by now see that for the time being, a restricted class of transseries will certainly do for physics.

The notation we will use here is to some extent borrowed from the theory of transseries but adapted in such a way as to suit our needs best. Let

$$\mathfrak{M} = \{\mathfrak{m}_1, ..., \mathfrak{m}_n, \mathfrak{m}_1^{-1}, ..., \mathfrak{m}_n^{-1}\} \tag{7.1.4}$$

be a finite set of symbols. We consider the set $\mathbb{R}[[\mathfrak{M}]]$ of formal transseries

$$f(\mathfrak{m}) = \sum_{l \in \mathbb{Z}^n} f_l \mathfrak{m}^l, \tag{7.1.5}$$

where $l = (l_1, ..., l_n) \in \mathbb{Z}^n$ is a multiindex and $\mathfrak{m}^l = \mathfrak{m}_1^{l_1} ... \mathfrak{m}_n^{l_n}$ is a formal product of elements in $\mathfrak{M}$, the transmonomial of order $l \in \mathbb{Z}^n$.

7.1.2. Transseries in quantum theory. In our exposition, however, we shall be content with $n = 3$ and the symbols

$$\mathfrak{m}_1 = z^a e^{-b/z} \ , \quad \mathfrak{m}_2 = z \ , \quad \mathfrak{m}_3 = \log(cz) \qquad (a, b, c \in \mathbb{R}) \tag{7.1.6}$$

and their inverses. These symbols will serve as basic building blocks of our transmonomials in this work. As alluded to in the introduction of this chapter, our choice of transmonomials is motivated by the tentative conjecture that the observables of a renormalisable QFT, seen as a function of the coupling $z \in \mathbb{C}$, belong to a class of functions known as *resurgent functions* which enjoy transseries representations of this form, hence called *resurgent transseries* [**DunUen14**].

Let us consider an example from quantum mechanics: the ground state energy of a quantum particle in a double-well potential

$$E(g) = \underbrace{\sum_{m \geq 0} c_{(0,m,0)} g^m}_{\text{perturbative part}} + \underbrace{\sum_{l_1, l_3 \geq 1} \sum_{l_2 \geq 0} \left(\frac{e^{-S/g}}{\sqrt{g}} \right)^{l_1} g^{l_2} \left\{ c_k^+ [\log(-2/g)]^{l_3} + c_k^- [\log(+2/g)]^{l_3} \right\}}_{\text{non-perturbative part}}, \tag{7.1.7}$$

treated as a 'perturbed' single well[2]. It has been obtained by using the WKB approach [**JenZin04**]. What we certainly learn from this expansion is that a conventional power series in the coupling g is not capable of capturing the whole physics of the double-well system: perturbation theory is bound to be blind to the 'flat sector', ie the non-perturbative part.

In particular, the 'log corrections' in (7.1.7) indicate that the perturbative part is *not Borel-summable* which leads to what is known as the *Stokes effect.* And the situation turns out to be no different in toy model QFTs like, for example, the $\mathbb{CP}^{N-1}$ model investigated in [**DunUen12,**

[1] The reader may ponder over this terminology.

[2] S is the 'instanton action' and g a parameter for which the double well becomes a single well in the limit $g \to 0$.

DunUen13]. This gives us enough reason to assume the same about renormalisable QFTs and carry out an investigation based on this assumption.

The point about *resurgence*, as developed in *Écalle's calculus of resurgent functions* [**Eca81**], is that *transseries are the suitable formal expansions to be used for Borel summation.* Because a suitably altered version of the Borel-Laplace transform can in this case reconstruct the function unambiguously, this introduces a new form of Borel summability, in some places referred to as *Écalle-Borel summability*[3].

Because we do not make use of Borel summation and resurgence theory in this work, we will not elaborate on these issues but rather take as a starting point the assumption that the *physical observables* of a renormalisable QFT have a transseries representation using only the three monomials in (7.1.6), more precisely:

ASSUMPTION 7.1 (Convenient working assumption). *The anomalous dimension $\gamma_1(z)$ and all higher RG functions $\gamma_2(z), \gamma_3(z), \ldots$ of a renormalisable QFT with a single coupling parameter z as introduced in Section 5.4 have resurgent transseries representations with transmonomials of the form*

$$\mathfrak{m}_1 = z^a e^{-b/z} \ , \quad \mathfrak{m}_2 = z \ , \quad \mathfrak{m}_3 = \log(cz). \tag{7.1.8}$$

The choice of parameters $a, b, c \in \mathbb{R}$ depends on the theory in question. For convenience, we set $a = 0, b = c = 1$ in this work.

The the choice of $b, c \in \mathbb{R}$ will not have any effect on the results presented in this chapter. However, changing the parameter $a \in \mathbb{R}$ will lead to different results, but only in minor detail, so that the overall messages conveyed remain the same.

For a thorough treatment of resurgence theory, we refer to the excellent mathematical introductions [**Sa07, Sa14**] and, for applications in physics, we recommend [**Dori14, Mar14**], written from the physicist's viewpoint.

Note that (7.1.7) does *not* imply that one can write the function $E(g)$ as a sum of the perturbative and the non-perturbative part. We have argued against the possibility of doing so already in Section 6.2. The expansion makes only sense as a whole: taking the Borel-Laplace transform of both pieces separately makes no sense; the perturbative part *alone* is in general not Borel-summable on account of the function's poles on the half-line $\mathbb{R}_+ \subset \mathbb{C}$ in the Borel plane.

7.1.3. Differential algebra structure. From now on, $\mathfrak{M}$ will be the set of elementary transmonomials defined by (7.1.4) and (7.1.6). The set of transseries $\mathbb{R}[[\mathfrak{M}]]$ is naturally an algebra, where the product is canonical and given by

$$f(\mathfrak{m})g(\mathfrak{m}) = (\sum_{l\in\mathbb{Z}^3} f_l \mathfrak{m}^l)(\sum_{k\in\mathbb{Z}^3} g_k \mathfrak{m}^k) = \sum_{l\in I} (f_. * g_.)_l \mathfrak{m}^l \tag{7.1.9}$$

and whose coefficients are computed through the convolution product

$$(f_. * g_.)_l := \sum_{l'+l''=l} f_{l'} g_{l''} \tag{7.1.10}$$

which involves a (finite) triple sum. Next, we introduce a *derivation* D on $\mathbb{R}[[\mathfrak{M}]]$ by setting[4]

$$\mathsf{D}(\mathfrak{m}_1) := \mathfrak{m}_1\mathfrak{m}_2^{-1} \ , \quad \mathsf{D}(\mathfrak{m}_2) := \mathfrak{m}_2 \ , \quad \mathsf{D}(\mathfrak{m}_2^{-1}) := -\mathfrak{m}_2^{-1} \ , \quad \mathsf{D}(\mathfrak{m}_3) := 1, \ \ldots \tag{7.1.11}$$

and so on, which, in essence, is nothing but $\mathsf{D} = z\partial_z$ acting on our transmonomials when taken seriously as functions of $z \in \mathbb{C}$.

The derivation D acts on a generic transmonomials $\mathfrak{m}^l = \mathfrak{m}_1^{l_1}\mathfrak{m}_2^{l_2}\mathfrak{m}_3^{l_3}$ according to

$$\mathsf{D}(\mathfrak{m}^l) = \mathsf{D}(\mathfrak{m}_1^{l_1}\mathfrak{m}_2^{l_2}\mathfrak{m}_3^{l_3}) = l_1\mathfrak{m}_1^{l_1}\mathfrak{m}_2^{l_2-1}\mathfrak{m}_3^{l_3} + l_2\mathfrak{m}_1^{l_1}\mathfrak{m}_2^{l_2}\mathfrak{m}_3^{l_3} + l_3\mathfrak{m}_1^{l_1}\mathfrak{m}_2^{l_2}\mathfrak{m}_3^{l_3-1} \tag{7.1.12}$$

[3]Or alternatively with both names swopped.

[4]Readers not acquainted with derivations should pause and first read Appendix Section A.7.

and gives rise to a differential structure on the algebra $\mathbb{R}[[\mathfrak{M}]]$, that is, the pair $(\mathbb{R}[[\mathfrak{M}]], \mathsf{D})$ is a differential algebra (an algebra equipped with a derivation).

To see explicitly how the derivation $\mathsf{D}\colon \mathbb{R}[[\mathfrak{M}]] \to \mathbb{R}[[\mathfrak{M}]]$ acts on transseries and, in particular, how this affects the coefficients, we use (7.1.12) and get

$$\mathsf{D}f(\mathfrak{m}) = \sum_{l\in\mathbb{Z}^3} \left[l_1 f_{\tau_2(l)} + l_2 f_l + (l_3+1) f_{\tau_3(l)} \right] \mathfrak{m}^l \ , \tag{7.1.13}$$

where the notation $\tau_2(l) := (l_1, l_2+1, l_3)$ and $\tau_3(l) := (l_1, l_2, l_3+1)$ is handy to account for the necessary index shift. Note that only the middle term would appear in the case of a standard power series, by which we mean a purely perturbative series $f(\mathfrak{m})$ whose coefficients $[\mathfrak{m}^l]f(\mathfrak{m}) = f_l$ satisfy $f_{(l_1,l_2,l_3)} = 0$ whenever $l_1 \neq 0$ or $l_3 \neq 0$.

7.2. RG transseries recursion as a discrete dynamical system

The reason we have introduced the differential structure is that we aim at investigating the RG function flow $\gamma_1 \longrightarrow \gamma_2 \longrightarrow \gamma_3 \longrightarrow ...$ represented as a discrete flow in $\mathbb{R}[[\mathfrak{M}]]$.

7.2.1. Transseries of RG functions. To this end, we consider again the sequence of functions $\gamma_1(z), \gamma_2(z), ...$ of the RG recursion

$$\gamma_n(z) = \frac{1}{n}\gamma_1(z)(sz\partial_z - 1)\gamma_{n-1}(z), \qquad (n \geq 2) \tag{7.2.1}$$

with single coupling $z \in \mathbb{C}$ and parameter $s = s_r \geq 0$ (see Section 5.4). We write the transseries representation $\widetilde{\gamma}_n(\mathfrak{m})$ of the n-th RG function $\gamma_n(z)$ in the form

$$\widetilde{\gamma}_n(\mathfrak{m}) = \sum_{l\in I} (\widetilde{\gamma}_n)_l \ \mathfrak{m}^l, \tag{7.2.2}$$

where $(\widetilde{\gamma}_n)_l = [\mathfrak{m}^l]\widetilde{\gamma}_n(\mathfrak{m})$ is the coefficient with triple index $l = (l_1, l_2, l_3) \in I := \mathbb{Z}^3$. We want the distinction between the function and its transseries to be reflected in the notation and adopt the convention of resurgence theorists by putting a tilde on top of the function's symbol whenever we mean its transseries [**Sa07**]. We write $\widetilde{\gamma}_n(\mathfrak{m}) \sim \gamma_n(z)$ to state this connection between the transseries $\widetilde{\gamma}_n(\mathfrak{m})$ and the resurgent function $\gamma_n(z)$ it purports to represent. We will refer to the transseries in $\widetilde{\gamma}_1(\mathfrak{m}), \widetilde{\gamma}_2(\mathfrak{m}), ...$ as *RG transseries* and its coefficients as *trans coefficients*[5]. In terms of transseries, the RG recursion reads

$$\widetilde{\gamma}_n(\mathfrak{m}) = \frac{1}{n}\widetilde{\gamma}_1(\mathfrak{m})(s\mathsf{D} - 1)\widetilde{\gamma}_{n-1}(\mathfrak{m}) = \frac{1}{n}\mathsf{R}_\theta(\widetilde{\gamma}_{n-1}(\mathfrak{m})), \tag{7.2.3}$$

where $\mathsf{R}_\theta := \theta(\mathfrak{m})(s\mathsf{D} - 1)$ is the RG recursion operator for a given anomalous dimension $\theta(\mathfrak{m}) := \widetilde{\gamma}_1(\mathfrak{m})$. The notation becomes somewhat neater, if we use $\theta_n(\mathfrak{m}) := n!\widetilde{\gamma}_n(\mathfrak{m})$ for all $n \geq 1$ so that

$$\theta_n(\mathfrak{m}) = \theta(\mathfrak{m})(s\mathsf{D} - 1)\theta_{n-1}(\mathfrak{m}) = \mathsf{R}_\theta^{n-1}(\theta_1(\mathfrak{m})), \tag{7.2.4}$$

where, of course $\theta_1(\mathfrak{m}) = \theta(\mathfrak{m})$ and $\mathsf{R}_\theta^m = \mathsf{R}_\theta \circ ... \circ \mathsf{R}_\theta$ is the m-fold composition of R_θ.

7.2.2. Discrete dynamical system. As alluded to above, this recursion gives rise to a discrete flow in the transseries algebra $\mathbb{R}[[\mathfrak{M}]]$. It follows the discrete-time evolution

$$\theta_{n+1}(\mathfrak{m}) = \mathsf{F}_\theta(n, \theta_1(\mathfrak{m})) \qquad \text{(RG transseries flow)}, \tag{7.2.5}$$

with discrete flow map $\mathsf{F}_\theta\colon \mathbb{N}_0 \times \mathbb{R}[[\mathfrak{M}]] \to \mathbb{R}[[\mathfrak{M}]]$ defined by $\mathsf{F}_\theta(n, \cdot) := \mathsf{R}_\theta^n$ for each anomalous dimension $\theta(\mathfrak{m}) = \widetilde{\gamma}_1(\mathfrak{m}) \in \mathbb{R}[[\mathfrak{M}]]$.

[5]As in 'Taylor coefficients'.

This makes for a family of *discrete dynamical systems* $\{(\mathsf{F}_\theta, \mathcal{X}, \mathcal{T}) : \theta \in \mathbb{R}[[\mathfrak{M}]]\}$, indexed by a candidate for the anomalous dimension $\theta(\mathfrak{m})$ and each consisting of the *state space* $\mathcal{X} = \mathbb{R}[[\mathfrak{M}]]$, the discrete time set $\mathcal{T} = \mathbb{N}_0$ and the flow map F_θ such that

$$\mathsf{F}_\theta(0, \cdot) = \mathrm{id}_{\mathcal{X}} \ , \qquad\qquad \mathsf{F}_\theta(n, \mathsf{F}_\theta(m, \cdot)) = \mathsf{F}_\theta(n+m, \cdot) \tag{7.2.6}$$

for all $n, m \in \mathcal{T}$. These two properties elevate the triple $(\mathsf{F}_\theta, \mathcal{X}, \mathcal{T})$ to a discrete dynamical system[6]. We will in the following happily switch between both notations $\theta_n(\mathfrak{m}) \leftrightarrow n!\widetilde{\gamma}_n(\mathfrak{m})$, as we believe there should be no serious potential for confusion.

However, although we may take any transseries $f(\mathfrak{m}) \in \mathbb{R}[[\mathfrak{m}]]$ as initial value and study its *orbit*

$$\mathsf{Orb}(f) = \{\mathsf{F}_\theta(n, f(\mathfrak{m})) : n \in \mathcal{T}\} \subset \mathbb{R}[[\mathfrak{M}]], \tag{7.2.7}$$

QFT asks explicitly for the orbit of the anomalous dimension $\theta(\mathfrak{m}) = \widetilde{\gamma}_1(\mathfrak{m})$ to which the index of the flow (map) F_θ refers, after all. In this view, the task posed by QFT is a hard one: find the initial transseries $\theta(\mathfrak{m}) \in \mathbb{R}[[\mathfrak{M}]]$ such that its orbit obeys the corresponding DSE as a condition. We will discuss this aspect at the end of this chapter in Section 7.5, where we treat the RG recursion in tandem with the DSE in the transseries setting.

7.3. Bigraded algebra of coefficients and homogeneous transseries

Note what the RG recursion tells us about the RG transseries: the trans coefficients of $\theta_n(\mathfrak{m})$ can be computed from those of the anomalous dimension $\theta(\mathfrak{m}) = \theta_1(\mathfrak{m})$, by virtue of the flow (7.2.5). The goal we set ourselves is now to monitor the flow of information between the different parts of this transseries and those of the higher RG transseries $\theta_n(\mathfrak{m})$. Since the coefficients of $\theta_n(\mathfrak{m})$ are just real numbers, the RG transseries flow is too oblivious for us to meet this goal.

7.3.1. Graded algebra of coefficients. It therefore makes sense to rephrase this RG transseries recursion in a more abstract setting. To this end, let $\mathscr{G} = \{c_l : l \in \mathbb{N}_0^3\}$ be a set of objects indexed by a triple index and let furthermore

$$\mathscr{A} := \mathbb{Q}[\mathscr{G}] = \mathbb{Q}[c_l : l \in \mathbb{N}_0^3] \tag{7.3.1}$$

be their freely generated commutative polynomial algebra over the rationals. Suppose we have an algebra morphism $\Phi \colon \mathscr{A} \to \mathbb{R}$ such that

$$\Phi(c_l) = (\widetilde{\gamma}_1)_l \qquad\qquad \forall l \in \mathbb{N}_0^3, \tag{7.3.2}$$

ie Φ maps every generator to the corresponding trans coefficient of the anomalous dimension $\widetilde{\gamma}_1(\mathfrak{m})$. This means that the transseries

$$\overline{\gamma}_1(\mathfrak{m}) := \sum_{l \in I} c_l \mathfrak{m}^l \in \mathscr{A}[[\mathfrak{M}]] \tag{7.3.3}$$

is related to the transseries of the anomalous dimension by $\widetilde{\gamma}_1(\mathfrak{m}) = \sum_{l\in I} \Phi(c_l)\mathfrak{m}^l$, where we set $c_l = 0$ for $l \notin \mathbb{N}_0^3$ because we want to keep $I = \mathbb{Z}^3$ as our summation index set. Note that also for physical reasons, we set $c_{(0,0,0)} = 0$ on account of $\gamma_n(0^+) = 0$ and $c_{(k,v,u)} = 0$ whenever $u \geq k$, because no such high-power log corrections are needed to compensate for the Stokes effect [**DunUen12**].

We will from now on place an overline on top of the corresponding symbol as in (7.3.3) to signify the difference between the transseries with real and the one with abstract coefficients in $\mathscr{A}$, ie

$$\Phi(\overline{\gamma}_n(\mathfrak{m})) = \widetilde{\gamma}_n(\mathfrak{m}) \ , \qquad\qquad \Phi(\overline{\theta}_n(\mathfrak{m})) = \theta_n(\mathfrak{m}) \tag{7.3.4}$$

[6]See any textbook on dynamical systems, eg [**Te12**] or [**Jo08**].

There is a canonical grading given by the number of generators that a product of elements from $\mathscr{G}$ exhibits:

Definition 7.2 (Length grading). *The derivation Y_0 on $\mathscr{A}$ defined by $Y_0(c_l) = c_l$ for any generator $c_l \in \mathscr{G}$ gives rise to a grading $\mathscr{A} = \bigoplus_{n\geq 0} \mathscr{G}_n$, where $x \in \mathscr{G}_n :\Leftrightarrow Y_0(x) = nx$ and $\mathscr{G}_0 = \mathbb{Q}1_{\mathscr{A}}$.*

This derivation just counts the number of generators in a product, ie its length if seen as a word. For example, for the product of two arbitrary generators, we have: $Y_0(c_l c_k) = Y_0(c_l)c_k + c_l Y_0(c_k) = 2c_l c_k$. This grading will be necessary for a thorough characterisation later on.

7.3.2. Bigrading. We define the two additional derivations $Y_1, Y_2 \colon \mathscr{A} \to \mathscr{A}$ by setting

$$Y_j(c_{(l_1,l_2,l_3)}) := l_j c_{(l_1,l_2,l_3)} \qquad (j = 1, 2) \tag{7.3.5}$$

for a generator $c_l = c_{(l_1,l_2,l_3)} \in \mathscr{G}$. In fact, these derivations give rise to a *bigrading*[7]:

$$\mathscr{A} = \bigoplus_{(u,v)\in\mathbb{N}_0^2} \mathscr{A}_{(u,v)}, \tag{7.3.6}$$

that is, $a \in \mathscr{A}_{(u,v)} :\Leftrightarrow Y_1(a) = ua$ and $Y_2(a) = va$. The subspace $\mathscr{A}_{(0,0)} = \mathbb{Q}1_{\mathscr{A}}$ is the kernel of both derivations. Note that the subspaces of this bigrading are finite-dimensional due to $c_{(0,0,0)} = 0$.

We shall refer to the associated derivations Y_1 and Y_2 as *instanton* and *loop grading operator*, respectively. The motivation for this denomination is that the transmonomial $\mathfrak{m}_1 = e^{-1/z}$ is associated to *instantons*, at least in asymptotically free theories [**DunUen12**].

7.3.3. Instanton and loop grading. Let us introduce two coarser gradings which are naturally implied by the bigrading, namely

$$\begin{aligned} \mathscr{M}_m &:= \bigoplus_{v\in\mathbb{N}_0} \mathscr{A}_{(m,v)} && \text{(instanton grading)} \\ \mathscr{L}_m &:= \bigoplus_{v\in\mathbb{N}_0} \mathscr{A}_{(v,m)} && \text{(loop grading)}, \end{aligned} \tag{7.3.7}$$

where we call the subspaces $\mathscr{M}_m \subset \mathscr{A}$ of the instanton grading *instanton sectors* of the coefficient algebra $\mathscr{A}$. Here is some more useful algebraic terminology that we will draw on subsequently:

Definition 7.3 (Gradings). *We say that $x \in \mathscr{A}$ is* homogeneous of degree $m \in \mathbb{N}_0$ *with respect to*

(i) *the instanton grading (or instanton-homogeneous), if $x \in \mathscr{M}_m$ ($:\Leftrightarrow Y_1(x) = mx$),*
(ii) *the loop grading (or loop-homogeneous), if $x \in \mathscr{L}_m$ ($:\Leftrightarrow Y_2(x) = mx$).*

For example, take $x = \alpha c_{(1,2,1)}c_{(1,1,0)} + \beta c_{(1,2,0)}c_{(t,5,1)} \in \mathscr{A}$ with $\alpha, \beta \in \mathbb{Q}$. This element is homogeneous of degree 2 with respect to the instanton grading only if $t = 1$ (first index is the instanton index),

$$Y_1(\alpha c_{(1,2,1)}c_{(1,1,0)} + \beta c_{(1,2,0)}c_{(t,5,1)}) = 2\alpha c_{(1,2,1)}c_{(1,1,0)} + (1+t)\beta c_{(1,2,0)}c_{(t,5,1)} \tag{7.3.8}$$

but not homogeneous regarding the loop grading (second index is the loop index):

$$Y_2(\alpha c_{(1,2,1)}c_{(1,1,0)} + \beta c_{(1,2,0)}c_{(t,5,1)}) = 3\alpha c_{(1,2,1)}c_{(1,1,0)} + 7\beta c_{(1,2,0)}c_{(t,5,1)}, \tag{7.3.9}$$

whatever value t assumes.

[7]See Appendix Section A.7 for a concise introduction to graded algebras.

7.3.4. Homogeneous transseries. These gradings, odd as this approach may strike the reader, will in fact help us to see what is going on as we step along the orbit of the RG transseries flow (7.2.5). We will elaborate on these issues in the next section. To be prepared for it, however, we need some more (straightforward) terminology, at the price of making the exposition even more idiosyncratic.

DEFINITION 7.4 (Homogeneity). *We call a transseries* $f(\mathfrak{m}) = \sum_{l\in I} f_l \mathfrak{m}^l \in \mathscr{A}[[\mathfrak{M}]]$

(i) homogeneous *with respect to the instanton grading (or instanton-homogeneous), if*

$$[\mathfrak{m}^l]f(\mathfrak{m}) = f_l \in \mathscr{M}_{l_1} \qquad \forall l = (l_1, l_2, l_3) \in I, \tag{7.3.10}$$

(ii) homogeneous *with respect to the loop grading (or loop-homogeneous), if*

$$[\mathfrak{m}^l]f(\mathfrak{m}) = f_l \in \mathscr{M}_{l_2} \qquad \forall l = (l_1, l_2, l_3) \in I. \tag{7.3.11}$$

We denote the subspaces of the corresponding homogeneous transseries by

$$\begin{aligned} \mathbb{T}_1(\mathfrak{M}) &:= \left\{ f(\mathfrak{m}) \in \mathscr{A}[[\mathfrak{M}]] : [\mathfrak{m}^{(l_1,l_2,l_3)}]f(\mathfrak{m}) \in \mathscr{M}_{l_1}, \forall l \in I \right\} \subset \mathscr{A}[[\mathfrak{M}]], \\ \mathbb{T}_2(\mathfrak{M}) &:= \left\{ f(\mathfrak{m}) \in \mathscr{A}[[\mathfrak{M}]] : [\mathfrak{m}^{(l_1,l_2,l_3)}]f(\mathfrak{m}) \in \mathscr{L}_{l_2}, \forall l \in I \right\} \subset \mathscr{A}[[\mathfrak{M}]]. \end{aligned} \tag{7.3.12}$$

This means that the degree of homogeneity of all coefficients corresponds exactly to the associated monomial in the transseries. Another way of writing the sets of these transseries is therefore

$$\begin{aligned} \mathbb{T}_1(\mathfrak{M}) &= \sum_{l\in I} \mathscr{M}_{l_1} \mathfrak{m}^l && \text{(instanton-homogeneous transseries)} \\ \mathbb{T}_2(\mathfrak{M}) &= \sum_{l\in I} \mathscr{L}_{l_2} \mathfrak{m}^l && \text{(loop-homogeneous transseries).} \end{aligned} \tag{7.3.13}$$

Apart from the fact that the zero transseries is homogeneous with respect to both gradings, note what a distinguished class such transseries are in $\mathscr{A}[[\mathfrak{M}]]$: pick any transseries $f(\mathfrak{m}) \in \mathscr{A}[[\mathfrak{M}]]$ and you will find that its coefficients may contain any elements in $\mathscr{A}$, that is, the summation index $l \in I$ has in general nothing to do with the indices of the elements from the generator set $\mathscr{G}$. Here is a nice little and straightforward

LEMMA 7.5 (Homogeneous subalgebras). *The subspaces* $\mathbb{T}_j(\mathfrak{M}) \subset \mathscr{A}[[\mathfrak{M}]]$ *for* $j = 1, 2$ *are subalgebras and so is their intersection*

$$\mathbb{T}_1(\mathfrak{M}) \cap \mathbb{T}_2(\mathfrak{M}) \subset \mathscr{A}[[\mathfrak{M}]]. \tag{7.3.14}$$

Furthermore, $\mathbb{T}_1(\mathfrak{m})$ *is* D*-stable, ie* $\mathsf{D}(\mathbb{T}_1(\mathfrak{M})) \subset \mathbb{T}_1(\mathfrak{M})$*, while the same is not true for* $\mathbb{T}_2(\mathfrak{M})$*.*

PROOF. Take $f(\mathfrak{m}), g(\mathfrak{m}) \in \mathbb{T}_j(\mathfrak{M})$, then the first assertion follows from

$$Y_j((f. * g.)_l) = \sum_{l'+l''=l} [Y_j(f_{l'})g_{l''} + f_{l'}Y_j(g_{l''})] = \sum_{l'+l''=l} \underbrace{[l'_j + l''_j]}_{=l_j} f_{l'}g_{l''} = l_j(f. * g.)_l. \tag{7.3.15}$$

As regards the latter assertion, we recall (7.1.13) for $f(\mathfrak{m}) \in \mathbb{T}_1(\mathfrak{M})$, ie

$$\mathsf{D}f(\mathfrak{m}) = \sum_{l\in I} (\mathsf{D}f)_l \mathfrak{m}^l = \sum_{l\in I} \left[l_1 f_{\tau_2(l)} + l_2 f_l + (l_3+1) f_{\tau_3(l)} \right] \mathfrak{m}^l \ , \tag{7.3.16}$$

and see that $Y_1((\mathsf{D}f)_l) = [l_1^2 f_{\tau_2(l)} + l_1 l_2 f_l + l_1(l_3+1) f_{\tau_3(l)}] = l_1(\mathsf{D}f)_l$ while

$$Y_2((\mathsf{D}f)_l) = l_1(l_2+1) f_{\tau_2(l)} + l_2^2 f_l + l_2(l_3+1) f_{\tau_3(l)} = l_2(\mathsf{D}f)_l + l_1 f_{\tau_2(l)} \tag{7.3.17}$$

shows that the same cannot in general hold for loop-homogeneous transseries. $\square$

Definition 7.6 (Instanton sector). *Let $f(\mathfrak{m}) \in \mathscr{A}[[\mathfrak{M}]]$. We refer to the trans subseries given by*

$$[\mathfrak{m}_1^k]f(\mathfrak{m}) = \sum_{(u,v)\in\mathbb{Z}^2} f_{(k,u,v)}\mathfrak{m}_2^u\mathfrak{m}_3^v \qquad \textit{('k-th instanton sector')} \tag{7.3.18}$$

as the k-th instanton sector *of $f(\mathfrak{m})$.*

Note that the perturbative part of a transseries is its zero-th instanton sector and what (7.3.17) informs us about is that the perturbative sector, if loop-homogeneous, will keep this feature under the action of D: the reason is, the distortion of the coefficients in (7.3.16) trivialises to

$$(\mathsf{D}f)_{(0,m,0)} = 0 \cdot f_{\tau_2(0,m,0)} + mf_{(0,m,0)} + f_{\tau_3(0,m,0)} = mf_{(0,m,0)} + f_{(0,m,1)} = mf_{(0,m,0)} \tag{7.3.19}$$

for all $m \in \mathbb{N}_0$. Since the homogeneity-preserving properties of a map acting on homogeneous transseries will turn out to be a crucial aspect, we shall devote to it a

Definition 7.7. *A map $\mathsf{M}\colon \mathscr{A}[[\mathfrak{M}]] \to \mathscr{A}[[\mathfrak{M}]]$ is called* homogeneity-preserving with respect to the instanton (or loop grading)*, if $\mathbb{T}_1(\mathfrak{M})$ (or $\mathbb{T}_2(\mathfrak{M})$) is stable under the action of M, ie if*

$$\mathsf{M}(\mathbb{T}_j(\mathfrak{M})) \subset \mathbb{T}_j(\mathfrak{M}) \tag{7.3.20}$$

for $j = 1$ (or $j = 2$), respectively.

7.4. RG recursion: the nonperturbative draws on the perturbative

Recall that we write $\overline{\theta}(\mathfrak{m}) \in \mathscr{A}[[\mathfrak{M}]]$ to denote the transseries associated to $\theta(\mathfrak{m}) \in \mathbb{R}[[\mathfrak{M}]]$ through the algebra morphism $\Phi\colon \mathscr{A} \to \mathbb{R}$ by $\Phi(\overline{\theta}(\mathfrak{m})) = \theta(\mathfrak{m})$.

7.4.1. RG iteration on the transseries algebra. With all the above defintions at hand, we will find that an important conclusion about the discrete RG flow in $\mathscr{A}[[\mathfrak{M}]]$ now, in fact, falls into our lap:

Corollary 7.8 (Discrete RG flow). *The RG recursion operator*

$$\mathsf{R}_{\overline{\theta}} = \overline{\theta}(\mathfrak{m})(s\mathsf{D} - 1_{\mathscr{A}})\colon \mathscr{A}[[\mathfrak{m}]] \to \mathscr{A}[[\mathfrak{M}]] \tag{7.4.1}$$

is homogeneity-preserving with respect to the instanton grading, that is, the orbit of the anomalous dimension $\overline{\theta}(\mathfrak{m}) = \overline{\gamma}_1(\mathfrak{m})$

$$\mathsf{Orb}(\overline{\theta}) = \Big\{\mathsf{F}_{\overline{\theta}}(n, \overline{\theta}(\mathfrak{m})) = \mathsf{R}^n_{\overline{\theta}}(\overline{\theta}(\mathfrak{m})) : n \in \mathcal{T}\Big\} \subset \mathbb{T}_1(\mathfrak{M}). \tag{7.4.2}$$

habours instanton-homogeneous transeries only*!*

Proof. By definition $\overline{\theta} \in \mathbb{T}_1(\mathfrak{M})$. Lemma 7.5 ensures that the assertion is true because $\mathbb{T}_1(\mathfrak{M})$ is stable under the structures of the differential algebra $\mathscr{A}[[\mathfrak{M}]]$. □

The assertion means that $Y_1((\overline{\theta}_n)_{(k,u,v)}) = k(\overline{\theta}_n)_{(k,u,v)}$ all along the orbit, ie for all $n \geq 1$, irrespective of the other indices. This entails that all coefficients of the k-th instanton sector

$$[\mathfrak{m}_1^k]\overline{\gamma}_n(\mathfrak{m}) = \sum_{(u,v)\in\mathbb{Z}^2} (\overline{\gamma}_n)_{(k,u,v)}\mathfrak{m}_2^u\mathfrak{m}_3^v \tag{7.4.3}$$

of the RG transseries $\overline{\gamma}_n(\mathfrak{m}) \in \mathbb{T}_1(\mathfrak{M})$ lie in the k-th instanton sector $\mathscr{M}_k$, ie this subseries receives no data from higher instanton sectors $\bigoplus_{j>k}\mathscr{M}_j$. This is because a generator $c_l \in \mathscr{G}$ with $Y_1(c_l) = l_1 c_l$ such that $l_1 > k$ cannot be found in $\mathscr{M}_k$. In particular, the perturbative sector with instanton degree $k = 0$ is completely unaffected by the nonperturbative sectors with instanton degree $k \geq 1$, while the converse is wrong because all instanton sectors $\mathscr{M}_k$ contain perturbative elements.

For example, the trans coefficients of $\overline{\gamma}_2(\mathfrak{m})$ are given by

$$(\overline{\theta}_2)_l = 2(\overline{\gamma}_2)_l = (\overline{\gamma}_1(s\mathsf{D} - 1_{\mathscr{A}})\overline{\gamma}_1)_l = \sum_{l'+l''=l} c_{l'}\left[s l_1'' c_{\tau_2(l'')} + (s l_2'' - 1)c_{l''} + s(l_3''+1)c_{\tau_3(l'')}\right]. \tag{7.4.4}$$

Its first few coefficients of the first instanton sector read

$$\begin{aligned} &2(\overline{\gamma}_2)_{(1,0,0)} = 0\ , \qquad 2(\overline{\gamma}_2)_{(1,0,1)} = 0 \\ &2(\overline{\gamma}_2)_{(1,1,0)} = (s-2)c_{(1,0,0)}c_{(0,1,0)} + s c_{(0,1,0)}[c_{(1,1,0)} + c_{(1,0,1)}] \ \in \mathscr{A}_{(1,1)} \oplus \mathscr{A}_{(1,2)} \subset \mathscr{M}_1 \\ &2(\overline{\gamma}_2)_{(1,1,1)} = c_{(0,1,0)}[(s-1)c_{(1,0,1)} + s c_{(1,1,1)} + s c_{(1,0,1)}] \ \in \mathscr{A}_{(1,1)} \oplus \mathscr{A}_{(1,2)} \subset \mathscr{M}_1, \end{aligned} \tag{7.4.5}$$

where we have indicated which subspaces and instanton sectors of the bigrading (7.3.6) the nonvanishing coefficients lie in. This shows that $\overline{\gamma}_2(\mathfrak{m})$ is only homogeneous with respect to the instanton grading, but not the loop grading. The RG recursion does not preserve homogeneity with respect to this grading due to the presence of the derivation D in $\mathsf{R}_{\overline{\theta}}$, unless $s=0$ as in the case of the ladder and rainbow approximations, which we shall come back to in due course.

7.4.2. Subspaces of trans coefficients. Note that by definition, the coefficients of the transseries $\overline{\theta}(\mathfrak{m})$ are all homogeneous of length degree 1 and that the action of the derivation D does not change this. In contrast to the action of $\mathsf{R}_{\overline{\theta}}$, which increases it by precisely 1. A simple observation is that the coefficient $(\overline{\theta}_n)_l \in \mathscr{G}_n$ is a linear combination of products of n generators, true for all triple indices $l \in \mathbb{N}_0^3$, where this includes the trivial cases of vanishing coefficients.

However, let us come back to the computation in (7.4.5): it suggests that the number of subspaces of the RG transseries' coefficients can only grow along the RG orbits. The next assertion specifies how exactly this is happening.

Proposition 7.9 (Subspace of trans coefficients). *The l-th coefficient of $\overline{\gamma}_n(\mathfrak{m}) \in \mathbb{T}_1(\mathfrak{M})$ in the nonperturbative sector satisfies*

$$(\overline{\gamma}_n)_l = (\overline{\gamma}_n)_{(l_1,l_2,l_3)} \in \mathscr{G}_n \cap \bigoplus_{k=0}^{n-1} \mathscr{A}_{\tau_2^k(l_1,l_2)} = \mathscr{G}_n \cap \bigoplus_{k=0}^{n-1} \mathscr{A}_{(l_1,l_2+k)} \subset \mathscr{M}_{l_1}, \tag{7.4.6}$$

and $(\overline{\gamma}_n)_{(0,m,0)} \in \mathscr{G}_n \cap \mathscr{A}_{(0,m)}$ in the perturbative sector. Furthermore, we have $(\overline{\gamma}_n)_l = 0$ if $n > l_1 + l_2$.

Proof. We only prove it for the nonperturbative sector, the perturbative case is an altered, in actual fact trivialised version of the presented proof, as no index shift happens in this case. We proceed by induction. The start of the induction $n=1$ is trivial because $(\overline{\gamma}_1)_{(l_1,l_2,l_3)} = c_{(l_1,l_2,l_3)} \in \mathscr{A}_{(l_1,l_2)}$ by definition of the bigrading. We use the shorthand notation

$$\mathcal{Y}(f_l) := l_1 f_{\tau_2(l)} + l_2 f_l + (l_3+1) f_{\tau_3(l)} \tag{7.4.7}$$

for the action on the coefficients representing the derivation D in (7.1.13). The RG recursion entails

$$\begin{aligned} c_{l'}(s\mathcal{Y} - 1_{\mathscr{A}})(\overline{\gamma}_n)_{l''} &\in c_{l'}(s\mathcal{Y} - 1_{\mathscr{A}}) \bigoplus_{k=0}^{n-1} \mathscr{A}_{\tau_2^k(l_1'',l_2'')} \subset c_{l'} \bigoplus_{k=0}^{n-1} \left\{ \mathscr{A}_{\tau_2^k(l_1'',l_2'')} \oplus \mathscr{A}_{\tau_2^{k+1}(l_1'',l_2'')} \right\} \\ &\subset \bigoplus_{k=0}^{n-1} \left\{ \mathscr{A}_{\tau_2^k(l_1,l_2)} \oplus \mathscr{A}_{\tau_2^{k+1}(l_1,l_2)} \right\} = \bigoplus_{k=0}^{n} \mathscr{A}_{\tau_2^k(l_1,l_2)} \end{aligned} \tag{7.4.8}$$

As regards the last assertion, if $n = l_1 + l_2$, then by $(\overline{\gamma}_n)_l \in \mathscr{G}_n$, one has

$$(\gamma_n)_{(l_1,l_2,l_3)} = [h_0\ c_{(1,0,0)}^{l_1} + h_1\ c_{(1,0,1)}^{l_1}] c_{(0,1,0)}^{l_2} \tag{7.4.9}$$

with $h_0, h_1 \in \mathbb{Z}$ (whatever l_3 is, it must be $l_3 = l_1$ or $l_3 = 0$, otherwise the whole caboodle vanishes). When $n > l_1 + l_2$, we need another generator in the product which does not increase the instanton or the loop grading degree. This can only be $c_{(0,0,0)} = 0$. □

Some straightforward conclusions can be drawn now.

(C1) Perturbative sector: the result for instanton sector $k = 0$ is hardly surprising. Yet we note for the record what it means: the coefficients of the *perturbative sector* contain no *nonperturbative data* and, by being 'length-homogeneous', their form is

$$(\overline{\gamma}_n)_{(0,u,0)} = \sum_{u_1+\ldots+u_n=u} h_{u_1\ldots u_n}\; c_{(0,u_1,0)} \cdots c_{(0,u_n,0)} \in \mathscr{M}_0 \cap \mathscr{L}_u \tag{7.4.10}$$

with, in fact integer coefficients $h_{u_1\ldots u_n} \in \mathbb{Z}$. In particular, we have $(\overline{\gamma}_n)_{(0,n,0)} \propto c^n_{(0,1,0)}$, which is the coefficient of the first term in the perturbative part of the n-th RG transseries due to $(\overline{\gamma}_n)_{(0,u,0)} = 0$ if $u < n$.

(C2) Nonperturbative sector: the nonperturbative piece of the RG transseries is more interesting:

$$(\overline{\gamma}_n)_{(k,u,v)} \in \mathscr{G}_n \cap \{\mathscr{A}_{(k,u)} \oplus \ldots \oplus \mathscr{A}_{(k,u+n-1)}\} \subset \mathscr{M}_k \cap (\mathscr{L}_u \oplus \ldots \oplus \mathscr{L}_{u+n-1}). \tag{7.4.11}$$

This shows that the coefficients of the k-th *nonperturbative (instanton) sector* contain only data from instanton sectors $\mathscr{M}_j$ with $j \leq k$. But what sets the nonperturbative sector apart from the perturbative one is that higher loop-order data is needed to compute these coefficients. Technically, this arises from the index shift on the loop index and (7.4.7) clearly points at its instantonic origin: no loop index shift can occur at perturbative level, where $l_1 = 0$.

We can make this last statement more precise, to be shown next:

- the perturbative coefficient $(\overline{\gamma}_n)_{(0,u,0)}$ has data of loop order up to $\ell_* = u - n + 1$,
- the nonperturbative coefficient $(\overline{\gamma}_n)_{(k,u,v)}$ contains data from contributions of loop order up to $\ell_* = u$.

These higher loop-order contributions both come from the perturbative and the nonperturbative sector of the anomalous dimension $\overline{\gamma}_1(\mathfrak{m}) = \sum_{l\in I} c_l \mathfrak{m}^l$. The next lemma makes the above two points above clear.

LEMMA 7.10 (Loop-order data). *The subspace $\mathscr{G}_n \cap \mathscr{A}_{(k,u)}$ only contains data of loop order $\ell \leq \ell_* := u - n + 1$, that is, a generator of the form $c_{(\cdot,\ell,\cdot)} \in \mathscr{G}$ such that $\ell > \ell_*$ is nowhere to be found in this subspace*[8]*. Furthermore, $\mathscr{G}_n \cap \mathscr{A}_{(k,u)} = \{0\}$ if $n > k + u$.*

PROOF. The latter statement is true for the same reason why $(\overline{\gamma}_n)_l = 0$ if $n > l_1 + l_2$, as expounded in the proof of Proposition 7.9. Note that any element in the subspace $\mathscr{G}_n \cap \mathscr{A}_{(k,u)} \subset \mathscr{G}_n \cap \mathscr{L}_u$ is necessarily a linear combination of terms of the form

$$x = c^m_{(\cdot,1,\cdot)} \prod_{t=1}^{r} c^{j_t}_{(\cdot,n_t,\cdot)} \in \mathscr{G}_n \cap \mathscr{L}_u \tag{7.4.12}$$

where $m + j_1 n_1 + \ldots + j_r n_r = u$ due to being an element in $\mathscr{L}_u$, $m + r = n$ by consisting of a product of n generators and $n_t \geq 2$ for all $t \in \{1, \ldots, r\}$ (we have separated out 1-loop data). The case $r = 0$ (ie, $m = n = u$) trivially yields $\ell_* = n - u + 1 = 1$, ie only first-loop order data is encapsulated. Let $r \geq 1$, ie there is higher loop order data in x now. Let $w \in \{1, \ldots, r\}$ such

[8] This loose statement means: set $c_{(\cdot,\ell,\cdot)} = 0$ in any $x \in \mathscr{G}_n \cap \mathscr{A}_{(k,u)}$ to find $x|_{c_{(\cdot,\ell,\cdot)}=0} = x$.

that $n_w = \max\{n_1, ..., n_r\}$ is the highest loop order which data in x can possibly have. We find

$$\ell_* - n_w = u - n + 1 - n_w = \underbrace{j_1 n_1 + ... + j_r n_r - r}_{=u-n} + 1 - n_w$$

$$= \underbrace{(j_w - 1)n_w}_{\geq 0} + \underbrace{\sum_{t \neq w} j_t n_t - (r-1)}_{\geq 0} \geq 0. \tag{7.4.13}$$

Consequently, x has only data from loop-order $\ell = n_w \leq \ell_*$. □

The positive integer $\ell_* = u - n + 1$ therefore presents an upper bound for the loop order of the data stored in the elements of the subspace $\mathscr{G}_n \cap \mathscr{A}_{(k,u)} \subset \mathscr{G}_n \cap \mathscr{L}_u$.

7.4.3. Transseries of rainbow and ladder RG. Notice what happens in the rainbow and ladder case: although the assertion of Proposition 7.9 remains valid, the situation trivialises substantially: according to (5.4.37), the RG recursion operator is simply given by a multiplication opertor in $\mathscr{A}[[\mathfrak{M}]]$: $\mathsf{R}_{\overline{\theta}} = \overline{\theta}(\mathfrak{m}) = \overline{\theta}_1(\mathfrak{m})$ and the discrete RG flow reads

$$\overline{\theta}_n(\mathfrak{m}) = \mathsf{R}_{\overline{\theta}}(\overline{\theta}_{n-1}(\mathfrak{m})) = \mathsf{R}_{\overline{\theta}}^{n-1}(\overline{\theta}_1(\mathfrak{m})) = \overline{\theta}_1(\mathfrak{m})^n \quad \text{(rainbows/ladders)}, \tag{7.4.14}$$

and, as a consequence, we find

Proposition 7.11. *The RG transseries of the rainbow and ladder approximations have homogeneous coefficients with respect to the bigrading, ie* $(\overline{\gamma}_n)_{(l_1,l_2,l_3)} \in \mathscr{A}_{(l_1,l_2)}$ *for all* $l \in I$ *and all along the orbit for all* $n \geq 1$.

Proof. Obvious from $\overline{\gamma}_1(\mathfrak{m})^n = \sum_{l \in I} (c^{*n}_{\cdot})_l \ \mathfrak{m}^l = \sum_{l \in I} \left(\sum_{s_1 + ... + s_n = l} c_{s_1} ... c_{s_n} \right) \mathfrak{m}^l$. □

This information about the rainbow and ladder RG recursions, however, gives us no hint about the fact that the anomalous dimension in these approximations is simply an algebraic function with a trivial transseries, ie a purely perturbative transseries, as is obvious from the results (5.2.19, 5.2.24). But we shall see in the next section that when combined with the DSE, an elementary computation proves that it trivialises to a purely perturbative series.

7.4.4. RG-driven flow of data. To summarise the state of play so far, according to the RG recursion, and hence the Callan-Symanzik equation, we note for the record that,

- firstly, the perturbative sector is completely independent of the nonperturbative sector as expected,
- secondly, *both perturbative and nonperturbative sectors of the anomalous dimension inform all sectors including the nonperturbative sectors of the higher RG functions in such a way that lower instanton sectors pass on data to all higher ones but never the other way around.*

The flow of information as driven by the RG transseries recursion is depicted schematically in Fig.1, where the non-shaded boxes represent the sources of data.

One has to admit that the conservation of homogeneity with respect to the instanton grading along the RG orbit is somewhat clear by considering the behaviour of the flat exponential functions represented by the transmonomial $\mathfrak{m}_1 = e^{-1/z}$: differentiating any power k of the exponential does not change k which in turn tells one which instanton sectors the accompanying coefficient cannot come from, namely the ones of higher degree than k. Although this result is therefore not entirely unexpected, the level of detail regarding the loop grading provided by Proposition 7.9 and Lemma 7.10 is noteworthy.

Anyhow, the foregoing analysis has been of twofold purpose: it firstly serves as a preparatory exercise and secondly gives a detailed picture to be utilised in the next section, where the methods will eventually be vindicated to the full.

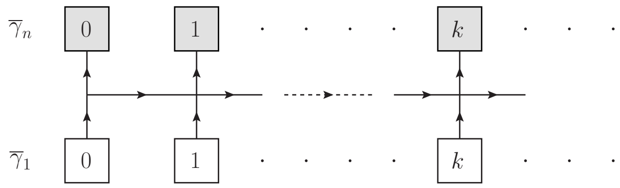


FIGURE 1. How the RG transseries recursion passes on perturbative and nonperturbative information from the anomalous dimension $\overline{\gamma}_1$ to the instanton sectors $0, 1, 2, ..., k$ of the n-th RG function $\overline{\gamma}_n$ $(n \geq 2)$. The non-shaded boxes, that is, the instanton sectors of the anomalous dimension, are the sources of information. The shaded boxes can only receive data.

7.5. How the perturbative determines the nonperturbative

Unfortunately, Figure 1 tells us nothing about a possible connection between the perturbative and the nonperturbative sectors of the anomalous dimension. Surely, by the way we set up this formalism, another result would have made little sense. We will in this section concern ourselves with what DSEs, or more precisely DSEs and RG recursion in combination, have to say about it.

Note that DSEs make no sense in the algebra $\mathscr{A}[[\mathfrak{M}]]$ of transseries with abstract coefficients because they establish relations between the different elements in $\mathscr{A}$, precisely what one does not have in a free commutative algebra. Take the Kilroy DSE (5.5.14) for the anomalous dimension

$$\gamma_1(z) = \frac{z}{2}\left[1 + \sum_{m\geq 1}\sum_{r\geq m}(-1)^r\ r! \sum_{r_1+...+r_m=r} \gamma_{r_1}(z)...\gamma_{r_m}(z)\right] \qquad \text{(Kilroy)}, \tag{7.5.1}$$

with Yukawa coupling z, for example. Recast in transseries language $\gamma_n(a) \sim \widetilde{\gamma}_n(\mathfrak{m})$ and combined with the RG recursion for transseries $n!\widetilde{\gamma}_n(\mathfrak{m}) = \mathsf{R}_\theta^{n-1}(\widetilde{\gamma}_1(\mathfrak{m}))$, this becomes a formidable equation in $\mathbb{R}[[\mathfrak{M}]]$:

$$\widetilde{\gamma}_1(\mathfrak{m}) = \frac{\mathfrak{m}_2}{2}\left[1 + \sum_{m\geq 1}\sum_{r\geq m}(-1)^r \sum_{r_1+...+r_m=r}\binom{r}{r_1, ..., r_m}\mathsf{R}_\theta^{r_1-1}(\widetilde{\gamma}_1(\mathfrak{m}))...\mathsf{R}_\theta^{r_m-1}(\widetilde{\gamma}_1(\mathfrak{m}))\right]. \tag{7.5.2}$$

If this identity were true in $\mathscr{A}[[\mathfrak{M}]]$ with coefficients in the free commutative algebra $\mathscr{A}$, then the generators $c_l \in \mathscr{G}$ would correspond to free parameters and this DSE had no information content whatsoever. Of course, this is not true and the obvious expedient is to read the relations between the various elements in $\mathscr{A}$ as an infinite system of coupled nonlinear difference equations to be solved in $\mathbb{R}$.

Before we study the Kilroy case, we briefly treat the linear approximations which have purely perturbative transseries and therefore make for a gentle and convenient start.

7.5.1. Ladders and rainbows revisited. We combine the RG recursion (5.4.37) and DSE (5.5.12) for the anomalous dimension of the ladder approximation to get the neat expression

$$\widetilde{\gamma}_1(\mathfrak{m}) = \mathfrak{m}_2\left[1 + \sum_{k=1}^{\infty}(-1)^k\mathsf{R}_\theta^{k-1}(\widetilde{\gamma}_1(\mathfrak{m}))\right] \qquad \text{(ladder RG + DSE)} \tag{7.5.3}$$

in the language of transseries. Or alternatively,

$$\left[1 + \frac{\mathfrak{m}_2}{1+\mathsf{R}_\theta}\right]\widetilde{\gamma}_1(\mathfrak{m}) = \mathfrak{m}_2, \tag{7.5.4}$$

which is impractical and useless yet nice to behold. The next assertion shows about the ladder and rainbow approximations what we already know: all RG functions have purely perturbative transseries because they are algebraic functions, see (5.2.19,5.2.24).

CLAIM 7.12 (Ladders and rainbows). *The combined message of both the RG recursion and the DSE of the ladder approximation stored in (7.5.3) is that* the perturbative dictates the nonperturbative sector to vanish. *As a result, all RG transseries of both the ladder approximation are purely perturbative and the same goes for the rainbow approximation.*

PROOF. On account of $\mathsf{R}_\theta = \widetilde{\gamma}_1(\mathfrak{m})$ being just a multiplication operator in $\mathbb{R}[[\mathfrak{M}]]$, we first employ (7.4.14) to rearrange (7.5.3) into

$$\widetilde{\gamma}(\mathfrak{m})^2 - \widetilde{\gamma}(\mathfrak{m}) = \mathfrak{m}_2, \tag{7.5.5}$$

which is (5.5.13) in terms of the anomalous dimension's transseries $\widetilde{\gamma}(\mathfrak{m}) = -\widetilde{\gamma}_1(\mathfrak{m}) \in \mathbb{R}[[\mathfrak{M}]]$. This equation is in fact an algebraic equation and the solution is an algebraic function, ie its transseries must be purely perturbative[9]. But let us see how this is revealed in a naive computation, not least because it makes for a nice exercise to get acquainted with transseries. We write the transseries as $\widetilde{\gamma}(\mathfrak{m}) = \sum_{l\in I} b_l \mathfrak{m}^l$ and get

$$\sum_{l\in I} [(b. * b.)_l - b_l]\ \mathfrak{m}^l = \mathfrak{m}_2 \tag{7.5.6}$$

with $(b.*b.)_l = \sum_{l'+l''=l} b_{l'} b_{l''}$. We are interested in the nonperturbative sector where $(b.*b.)_l = b_l$ and therefore start the computation with the first instanton sector whose coefficients, when subjected to this nonperturbative condition, ie (7.5.6), yield

$$b_{(1,t,s)} = 2\sum_{i=0}^{t}\sum_{j=0}^{s} b_{(1,i,j)} b_{(0,t-i,s-j)} = 2\sum_{i=0}^{t} b_{(1,i,s)} b_{(0,t-i,0)}, \tag{7.5.7}$$

where the last step is permitted on account of $b_{(0,t-i,s-j)} = 0$ if $j \neq s$. We first check the case $t = 0$, which gives $b_{(1,0,s)} = 2b_{(1,0,s)} b_{(0,0,0)}$. Here is where the perturbative sector tells both nonperturbative coefficients for $s = 0, 1$ to vanish: (7.5.5) requires $b_{(0,0,0)} \in \{0, 1\}$ for the perturbative coefficient. This enforces $b_{(1,0,s)} = 0$. Next, we do the induction step $t \to t + 1$ and find

$$b_{(1,t+1,s)} = 2\sum_{i=0}^{t+1} b_{(1,i,s)} b_{(0,t+1-i,0)} = 2b_{(1,t+1,s)} b_{(0,0,0)}, \tag{7.5.8}$$

where for the second equality we have made use of the assumption $b_{(1,i,s)} = 0$ for all $i \leq t$. Again, we find $b_{(1,t+1,s)} = 0$, again dictated by the perturbative coefficient. This entails that the whole first instanton sector is completely absent. Now assume that all instanton sectors of the transseries vanish up to the n-th. This statement means

$$b_{(n,t,s)} = \sum_{l=0}^{n}\sum_{i=0}^{t}\sum_{j=0}^{s} b_{(l,i,j)} b_{(n-l,t-i,s-j)} = 0 \tag{7.5.9}$$

for all $t, s \geq 0$. To see whether this is still true for $(n + 1)$, we have to scour

$$b_{(n+1,t,s)} = \sum_{l=0}^{n+1}\sum_{i=0}^{t}\sum_{j=0}^{s} b_{(l,i,j)} b_{(n+1-l,t-i,s-j)} \tag{7.5.10}$$

[9] The proof may therefore be finished at this stage.

for survivors. By assumption, the pieces of the outermost sum vanish for $l = 1, ..., n$, while the two remaining pieces for $l = 0$ and $l = n + 1$ yield the same expression and we arrive at

$$b_{(n+1,t,s)} = 2\sum_{i=0}^{t}\sum_{j=0}^{s} b_{(n+1,i,j)}b_{(0,t-i,s-j)} = 2\sum_{i=0}^{t} b_{(n+1,i,s)}b_{(0,t-i,0)} \tag{7.5.11}$$

as in (7.5.7) which vanishes again first for $t = 0$ and then by induction for all $t \geq 1$, just like we had in (7.5.7). □

To glimpse the launch of the cascade of collective vanishing, here are the first equations, as prescribed by (7.5.7)

$$\begin{aligned} &b_{(1,0,0)} = 2b_{(0,0,0)}b_{(1,0,0)}\ , \qquad b_{(1,0,1)} = 2b_{(0,0,0)}b_{(1,0,1)} \\ &b_{(1,1,0)} = 2b_{(0,0,0)}b_{(1,1,0)} + 2b_{(1,0,0)}b_{(0,1,0)} \\ &b_{(1,1,1)} = 2b_{(0,0,0)}b_{(1,1,1)} + 2b_{(0,1,0)}b_{(1,0,1)}. \end{aligned} \tag{7.5.12}$$

This shows nicely: whatever value the perturbative coefficient takes, ie $b_{(0,0,0)} \in \{0, 1\}$, this cascade cannot be stopped.

7.5.2. Kilroy flow. Let us now come to the first nontrivial case. The results about the discrete RG flow within the instanton-homogeneous subalgebra $\mathbb{T}_1(\mathfrak{M}) \subset \mathscr{A}[[\mathfrak{M}]]$ enable us now to see what is going on: the DSE can be interpreted as an *asymptotic constraint for the RG flow as a dynamical system.* First note that the operator family $\{\mathsf{Q}_m\}_{m\geq 1}$, given by

$$\mathsf{Q}_m(\overline{\theta}(\mathfrak{m})) := \sum_{r\geq m}(-1)^r \sum_{r_1+...+r_m=r} \binom{r}{r_1, ..., r_m} \prod_{j=1}^{m} \mathsf{R}_{\overline{\theta}}^{r_j-1}(\overline{\theta}(\mathfrak{m})) \tag{7.5.13}$$

is *homogeneity-preserving with respect to the instanton grading* on account of Corollary 7.8, which means that they have the instanton-homogeneous subalgebra as a common stable domain:

$$\mathsf{Q}_m(\mathbb{T}_1(\mathfrak{M})) \subset \mathbb{T}_1(\mathfrak{M}) \qquad \forall m \in \mathbb{N}. \tag{7.5.14}$$

The coefficients of this transseries can be characterised more precisely.

LEMMA 7.13. *For all $m \geq 1$ and $u, v \in \mathbb{N}_0$, one has*

$$[\mathfrak{m}^{(k,u,v)}]\mathsf{Q}_m(\overline{\theta}(\mathfrak{m})) \in \bigoplus_{j=0}^{j_*} \mathscr{A}_{(k,u+j)} \subset \mathscr{M}_k\ , \tag{7.5.15}$$

where $j_* = k + u - m$, in particular $[\mathfrak{m}^{(k,u,v)}]\mathsf{Q}_m(\overline{\theta}(\mathfrak{m})) = 0$ in case $m > k + u$.

PROOF. Consider

$$[\mathfrak{m}^{(k,u,v)}]\mathsf{Q}_m(\overline{\theta}(\mathfrak{m})) = \sum_{r\geq m}(-1)^r \sum_{r_1+...+r_m=r} \binom{r}{r_1, ..., r_m} [\mathfrak{m}^{(k,u,v)}] \prod_{j=1}^{m} \mathsf{R}_{\overline{\theta}}^{r_j-1}(\overline{\theta}(\mathfrak{m})). \tag{7.5.16}$$

First note that by Proposition 7.9, $[\mathfrak{m}^{(k,u,v)}]\mathsf{R}_{\overline{\theta}}^{n-1}(\overline{\theta}(\mathfrak{m})) = 0$ if $k + u < n$ and otherwise

$$[\mathfrak{m}^{(k,u,v)}]\mathsf{R}_{\overline{\theta}}^{n-1}(\overline{\theta}(\mathfrak{m})) \in \mathscr{G}_n \cap \bigoplus_{j=0}^{n-1} \mathscr{A}_{(k,u+j)}. \tag{7.5.17}$$

And hence by $[\mathfrak{m}^l]\{\mathsf{R}_{\overline{\theta}}^{n-1}(\overline{\theta}(\mathfrak{m}))\mathsf{R}_{\overline{\theta}}^{n'-1}(\overline{\theta}(\mathfrak{m}))\} = \sum_{l'+l''=l}[\mathfrak{m}^{l'}]\mathsf{R}_{\overline{\theta}}^{n-1}(\overline{\theta}(\mathfrak{m}))[\mathfrak{m}^{l''}]\mathsf{R}_{\overline{\theta}}^{n'-1}(\overline{\theta}(\mathfrak{m}))$, we find

$$[\mathfrak{m}^{(k,u,v)}]\{\mathsf{R}_{\overline{\theta}}^{n-1}(\overline{\theta}(\mathfrak{m}))\mathsf{R}_{\overline{\theta}}^{n'-1}(\overline{\theta}(\mathfrak{m}))\} \in \mathscr{G}_{n+n'} \cap \bigoplus_{j=0}^{n+n'-2} \mathscr{A}_{(k,u+j)} \tag{7.5.18}$$

for $n+n' \leq k+u$, otherwise the whole caboodle on the lhs vanishes and the statement is still true ('subspaces'). More generally, for $m < r = r_1 + ... + r_m$, we have

$$[\mathfrak{m}^{(k,u,v)}] \prod_{t=1}^{m} \mathsf{R}_{\overline{\theta}}^{r_t - 1}(\overline{\theta}(\mathfrak{m})) \in \mathscr{G}_r \cap \bigoplus_{j=0}^{r-m} \mathscr{A}_{(k,u+j)}, \tag{7.5.19}$$

whose lhs also vanishes if $r > k+u$. This means that the sum in (7.5.16) terminates at $r = k+u$ and the assertion follows. □

So we see that for all $m \geq 1$, the k-th instanton sector of the transseries $\mathsf{Q}_m(\overline{\theta}(\mathfrak{m})) \in \mathbb{T}_1(\mathfrak{M})$ has no nonperturbative information from higher instanton sectors $\bigoplus_{j>k} \mathscr{M}_j$ simply because this transseries is instanton-homogeneous.

As alluded to in Section 7.2, the task of finding the transseries for the anomalous dimension of the Kilroy approximation is now, in the language of dynamical systems, to find the right initial value $\theta_1(\mathfrak{m}) \in \mathbb{R}[[\mathfrak{M}]]$ such that its orbit under the RG flow

$$\theta_{n+1}(\mathfrak{m}) = \mathsf{F}_\theta(n, \theta_1(\mathfrak{m})) = \mathsf{R}_\theta^n(\theta_1(\mathfrak{m})) \tag{7.5.20}$$

satisfies the asymptotic constraint

$$\theta_1(\mathfrak{m}) = \lim_{N\to\infty} \frac{\mathfrak{m}_2}{2}[1 + \sum_{m=1}^{N} \mathsf{Q}_m(\theta_1(\mathfrak{m}))] \qquad \text{(asymptotic constraint)} \tag{7.5.21}$$

Proposition 7.14 (Kilroy transseries). *Let Assumption 7.1 be true and let $\theta_1(\mathfrak{m}) = \widetilde{\gamma}_1(\mathfrak{m})$ be a solution of the DSE (7.5.2), ie the 'Kilroy transseries'. Then its nonperturbative part is completely determined by its perturbative part. The recipe for computing all coefficients is then given by*

$$[\mathfrak{m}^{(k,u,v)}]\overline{\theta}_1(\mathfrak{m}) = [\mathfrak{m}^{(k,u,v)}] \frac{\mathfrak{m}_2}{2} \left[1 + \sum_{m=1}^{k+u} \mathsf{Q}_m(\overline{\theta}_1(\mathfrak{m}))\right] \tag{7.5.22}$$

plus an appropriate initial condition.

Proof. By Lemma 7.13 we find for fixed instanton and loop indices $k, u \in \mathbb{N}$, the finite system of nonlinear difference equations

$$[\mathfrak{m}^{(k,u,v)}]\overline{\theta}_1(\mathfrak{m}) = \lim_{N\to\infty} [\mathfrak{m}^{(k,u,v)}] \frac{\mathfrak{m}_2}{2} \left[1 + \sum_{m=1}^{N} \mathsf{Q}_m(\overline{\theta}_1(\mathfrak{m}))\right] = [\mathfrak{m}^{(k,u,v)}] \frac{\mathfrak{m}_2}{2} \left[1 + \sum_{m=1}^{k+u} \mathsf{Q}_m(\overline{\theta}_1(\mathfrak{m}))\right] \tag{7.5.23}$$

where the sum terminates on account of Lemma 7.13. This system is comprised of a finite number of equations, namely for $v \leq k$ and contains no elements from $\bigoplus_{j>k} \mathscr{M}_j$ and thus relies solely on lower instanton sectors, ie information in the lower pieces of the filtration, namely the subspace $\bigoplus_{j\leq k} \mathscr{M}_j$. □

Figure 2 shows schematically what this essentially means: the perturbative sector $\mathscr{M}_0$ of the anomalous dimension is the sole source of information and determines all of its own higher instanton sectors $\mathscr{M}_k$ for $k \geq 1$ and thereby all instanton sectors of the higher RG transseries.

The nonperturbative constraint (7.5.23) encodes, as we have explained in the introduction to this chapter, how to compute the nonperturbative coefficients from the perturbative ones. This is the precise sense in which the perturbative sector determines all nonperturbative sectors: the equations of instanton sector $k \geq 1$ have the coefficients from sectors $j \leq k$ as external parameters.

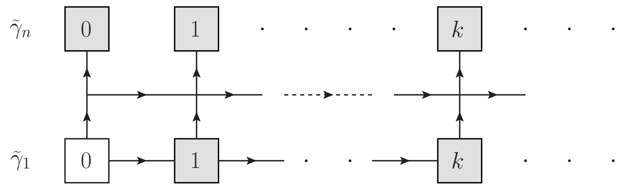


FIGURE 2. DSE and RG recursion for the Kilroy transseries team up to empower the perturbative sector of the anomalous dimension $\overline{\gamma}_1$ to determine its own nonperturbative sectors and thereby those of the n-th RG function $\overline{\gamma}_n$ ($n \geq 2$). The perturbative sector of the anomalous dimension, represented by the single non-shaded box in the lower left corner, is the only source of information.

7.5.3. Transseries in QED. In principle, this argument also applies to QED, for which the analogue of (7.5.2) reads in case of the anomalous dimension of the photon

$$\widetilde{\gamma}_1^{\sim}(\mathfrak{m}) = J(\mathfrak{m}_2) + \sum_{m\geq 1}\sum_{r\geq m} J_{m,r}(\mathfrak{m}_2) \sum_{r_1+\ldots+r_m=r} \binom{r}{r_1,\ldots,r_m} \prod_{j=1}^{m} \mathsf{R}_{\widetilde{\gamma}_1^{\sim}}^{r_j-1}(\widetilde{\gamma}_1^{\sim}(\mathfrak{m})), \tag{7.5.24}$$

which is simply (5.5.21) with the two 'skeleton' functions $J(\alpha)$ and $J_{m,r}(\alpha)$ introduced in (5.5.22). Yet there is a subtlety lurking in these functions. There are two scenarios:

(1) Both skeleton functions are Borel-summable and therefore need no nonperturbative completion, ie no additional transmonomials other than $\mathfrak{m}_2 = z$ are required in the transseries.
(2) This is not the case and at least one of them is not Borel-summable so that a nonperturbative completion is unavoidable.

The reader notices by now that (7.5.24) is an optimistic statement favouring the first scenario. In this case, the argument in Proposition 7.14 goes also through in QED. As far as Borel-summable series are concerned, there is a 'but': modern resurgence theory teaches us that there are conditions on such functions to allow for a unique reconstruction by means of Borel resummation. And a case where these conditions are not met has been encountered: latterly, [**GraMaZa15**] report on a perturbative genus expansion in string theory where this seems to occur due to poles off the real axis ('complex instantons').

Besides, one has to acknowledge that the second scenario is more likely on account of the rapidly growing number of Dyson-Schwinger skeletons. Nevertheless, this is easy to rectify by simply considering the nonperturbative completion of (7.5.24), ie

$$\widetilde{\gamma}_1^{\sim}(\mathfrak{m}) = \widetilde{J}(\mathfrak{m}) + \sum_{m\geq 1}\sum_{r\geq m} \widetilde{J}_{m,r}(\mathfrak{m}) \sum_{r_1+\ldots+r_m=r} \binom{r}{r_1,\ldots,r_m} \prod_{j=1}^{m} \mathsf{R}_{\widetilde{\gamma}_1^{\sim}}^{r_j-1}(\widetilde{\gamma}_1^{\sim}(\mathfrak{m})), \tag{7.5.25}$$

in which the skeleton functions are now represented by fully-fledged transseries. We can now again treat the rhs of this equation algebraically. This time however, we take the trans coefficients of $\widetilde{J}(\mathfrak{m})$ and $\widetilde{J}_{m,r}(\mathfrak{m})$ and add them as generators to the algebra $\mathscr{A}$ to obtain the 'skeleton-augmented' coefficient algebra $\mathscr{A}'$. Then, the rhs is clearly homogeneity preserving with respect to the instanton grading $k \mapsto \mathscr{M}'_k$ and we find

PROPOSITION 7.15 (Anomalous dimension of the photon). *If Assumption 7.1 is true and the photon transseries $\widetilde{\gamma}_1^{\sim}(\mathfrak{m})$ solves the DSE (7.5.25), then its perturbative part determines its nonperturbative part completely.*

Notice that we have phrased the two propositions 7.14 and 7.15 carefully within the transseries setting. The reason is twofold.

- No one yet knows the exact form of the anomalous dimension's transseries, let alone whether it is a resurgent function at all and how many poles it has on the Borel plane and where they are.
- The passage from the world of resurgent transseries to the corrosponding resurgent function is not only far from trivial but paved with ambiguities: each pole in the Borel plane poses a choice problem to which physics must provide the answer.

Nevertheless, in the light of the aforementioned recent auspicious developments, we do not expect the state of affairs to stagnate regarding these two points.

Additionally, we believe that the machinery developed in this chapter can be employed to investigate the possible form of the transseries: if a transseries ansatz does not work and refuses to cohere with a given Dyson-Schwinger equation by yielding insolvable algebraic nonlinear difference equations in $\mathbb{R}$, then we will have to try another one with different transmonomials.

Conclusion

We have reviewed Haag's theorem and some pertinent triviality results, scrutinising in particular the details of the proof of Haag's theorem and its provisions. The most salient provision turned out to be unitary equivalence between free and interacting quantum fields.

Because the theorem is independent of the dimension of Minkowski spacetime[10], it holds also for superrenormalisable quantum fields. A circumvention scheme in this context has been found by constructive field theorists: (super)renormalisation. And here, to the best of our knowledge, these authors do not claim unitary equivalence.

Although the case against quantum electrodynamics (and hence also quantum chromodynamics) is less clear due to a fundamental incompatibility with Wightman's framework, there is no doubt that the interaction picture can also not exist there.

Renormalisation against triviality. We have argued that

RENORMALISATION BYPASSES HAAG'S THEOREM

in all cases by effectively rendering the field intertwiner non-unitary. This cannot be proved due to the mathematical elusiveness of the involved renormalisation Z factors: per se only defined perturbatively to master their task of subtracting divergences, their nonperturbative status is totally obscure.

In particular, the wave-function renormalisation strikes us as a tenuously locked piece of the jigsaw puzzle that quantum field theory presents itself as. We contend that this 'constant' cannot convincingly establish the link between the spectral representation and Haag-Ruelle (or LSZ) scattering theory[11] with standard perturbative quantum field theory, the biggest and best-understood jigsaw chunk. Interpreting this renormalisation constant as a probability amplitude, as standard textbooks do, is something we would like to ask posterity's opinion about.

Another important theme are the canonical (anti)commutation relations for quantum fields. The germane triviality theorems we have discussed here suggest that these relations are incompatible with interactions in (flat) spacetimes of dimensions $d \geq 4$. With these results, we call into question the status of such relations in these spacetimes, in particular on the grounds that there is no reasonable analogue of the position operator in quantum field theory. This is intimately tied up with the Heisenberg uncertainty relations whose implementation is more or less obscure in QFT, a side issue we have only touched upon briefly and in passing.

Despite all these conceptual conundrums, once we run up against divergent integrals in the perturbative series of an n-point function of a quantum field, a change of perspective is inevitable. At this point, we switch to the combinatorial stance from which this ill-defined series re-emerges as a formal power series whose coefficients are formal pairs containing combinatorial information about distributions, encoded in Feynman diagrams. The task is then to get the mathematically well-defined game of Hopf-algebraic renormalisation started.

[10] Spacetime must have dimension $d \geq 2$, otherwise we would have either space or time, in particular would we have no boosts.

[11] The LSZ reduction formula, being very instructive for amputating Green's functions in momentum space, makes total sense nonetheless. We do not take a nihilist stance here.

In fact, one may adopt an even more radical attitude and say that formal power series with coefficients made up of formal pairs are simply the way quantum field theory speaks to us via perturbation theory, just as divergent asymptotic series in the context of nonlinear singular differential equations have data about the solution encapsulated, without making direct sense as functions. But we do not go that far.

We have seen that the virtue of the combinatorial approach lies in the generic way Dyson-Schwinger equations can be formulated for combinatorial power series, while the Callan-Symanzik equation arises naturally from the coproduct formula and the concept of Hopf algebra characters and their Lie generators.

Dyson-Schwinger equations in terms of Mellin transforms enable us to derive a system of formal equations for the anomalous dimension and all higher log-coefficient functions. Combined with the renormalisation group (RG) recursion, it is then in some cases possible to formulate a nonlinear ordinary differential equation for the anomalous dimension.

Photon equation & Landau pole. We have investigated one such equation in quantum electrodynamics ('photon equation') more thoroughly and found some criteria for the existence of a Landau pole. An exact toy model solution proved having a nonperturbative 'flat' contribution whose hampering impact on the beta function and the running coupling we have studied. In this model, it turned out that the flat contribution indeed hampers the growth of the beta function as compared to the result of first-order perturbation theory. We tentatively infer from this that

NONPERTURBATIVE FEATURES OF THE ANOMALOUS DIMENSION MATTER,

as they may have a decisive impact on the beta function and hence the running coupling in a more realistic real-world model.

Transseries & Dyson-Schwinger equations. On the assumtion that the anomalous dimension is a resurgent function, we have studied the RG recursion using a transseries ansatz. It turned out to be useful to think of this recursion as a discrete dynamical system with orbits in the set of transseries. The Dyson-Schwinger equation for the anomalous dimension then appeared in this view as an asymptotic constraint imposed on the orbit of the RG recursion. We have then proved that this constraint enforces nonlinear difference equations on the coefficients of the transseries in such a way that

THE PERTURBATIVE SECTOR DETERMINES THE NONPERTURBATIVE ONE COMPLETELY,

which we found for the anomalous dimension of the Yukawa fermion in the rainbow and ladder equations (trivial cases), the Kilroy equation (first nontrivial case) and finally for the anomalous dimension of the photon in quantum electrodynamics (least trivial).

The little machinery we have developed in this part of our work can in principle be employed to check various transseries ansätze and to see which one is working and satisfies the equations at hand, ie which one leads to solvable nonlinear difference equations for the coefficients of the resurgent transseries.

In an ideal course of future developments, we envision an explicit recursion formula for the trans coefficients of the anomalous dimension whose Borel-Écalle sum we can then plot on a computer screen.

APPENDIX A

Mathematical background material

This appendix chapter is a collection of mathematical definitions and some basic facts which the reader should be familiar with to be able to follow the deliberations of this work. Proofs are given only in some cases. The presentation is meant to be pedagogical except in Section A.1 on operator theory, which stands out in this respect: while we do not expect physicists to be familiar with Hopf algebras, we assume that this is the case with operator theory, as it is generally part of the physics curriculum.

In the following, we assume all vector spaces to be vector spaces over the field $\mathbb{F}$, either $\mathbb{R}$ or $\mathbb{C}$. In some cases $\mathbb{Q}$ or even simpler fields are sufficient[1]. The material collected in this part of the appendix is standard and can be found in classical as well as more recent books: on Hopf algebras, see [**Sw69, Ka12**], Birkhoff decompositions are for example described in [**Man04**].

A.1. Operators on Hilbert spaces

Hilbert spaces. A *pre-Hilbert space* is a vector space over $\mathbb{C}$ equipped with a scalar product $\langle \cdot, \cdot \rangle$ which is linear in the second and antilinear in the first slot[2]. A pre-Hilbert space is called *Hilbert space* X if it is complete with respect to its induced norm $||x|| := \sqrt{\langle x, x \rangle}$ for $x \in X$, ie any Cauchy sequence converges to an element in X in this norm. A subset $D \subset X$ is called *dense* (in X) if for any $x \in X$ and every $\epsilon > 0$, there is an element in $y \in D$ such that $||x - y|| < \epsilon$. A subset is called *subspace* if it is a Hilbert space. Let $I \subset \mathbb{R}$ be an index set. A set $\{e_\alpha : \alpha \in I\}$ of vectors in X is called *basis* of X if the set of their finite linear combinations is dense in X. If the index set I of the basis is countable, X is called *separable*.

Completion of pre-Hilbert spaces. Let X is a pre-Hilbert space and $\mathrm{CS}(X)$ denote the set of its Cauchy sequences. Note that for any Cauchy sequence $x = (x_n) \in \mathrm{CS}(X)$, the limit $\lim_{n\to\infty} ||x_n||$ exists since the sequence $(||x_n||)$ is Cauchy in $\mathbb{C}$ due to

$$| \, ||x_n|| - ||x_m|| \, | \leq ||x_n - x_m||. \tag{A.1.1}$$

We introduce an equivalence relation on $\mathrm{CS}(X)$: we say that $x, y \in \mathrm{CS}(X)$ are equivalent, in signs $x \sim y$, if $\lim_{n\to\infty} ||x_n - y_n|| = 0$. We write the equivalence class of $x \in \mathrm{CS}(X)$ as $[x] = \{x' \in \mathrm{CS}(X) : x' \sim x\}$. Then $\mathrm{CS}(X)/\sim$ is a vector space which we can endow with a scalar product given by

$$\langle [x], [y] \rangle_\sim := \lim_{n\to\infty} \langle x_n, y_n \rangle \tag{A.1.2}$$

for $[x], [y] \in \mathrm{CS}(X)/\sim$. This is well-defined: on account of

$$|\langle x_n, y_n \rangle - \langle x_m, y_m \rangle| = |\langle x_n, y_n - y_m \rangle + \langle x_n - x_m, y_m \rangle| \tag{A.1.3}$$

$$= ||x_n|| \, ||y_n - y_m|| + ||x_n - x_m|| \, ||y_m|| \tag{A.1.4}$$

the sequence $(\langle x_n, y_n \rangle)_n$ is Cauchy in $\mathbb{C}$. If we identify all elements of the pre-Hilbert space X with constant sequences in $\mathrm{CS}(X)$, then $\mathrm{CS}(X)/\sim$ is Hilbert with scalar product $\langle \cdot, \cdot \rangle_\sim$. This Hilbert space is called the *completion* of X.

[1] We require the field to have characteristic zero, though.

[2] Physics convention

Operators. Let A be a linear operator on a Hilbert space X. The linear subset $D(A) \subset X$ on whose elements it is defined is called its *domain*. A is said to *bounded* on $D(A)$ if there is a number $C > 0$ such that

$$||Ax|| \leq C||x|| \qquad \forall x \in D(A), \tag{A.1.5}$$

otherwise *unbounded*. We denote the set of bounded linear operators on X by $B(X)$. We shall drop 'linear' from now on since all operators will be linear in this appendix section. The *operator norm* on $B(X)$ is defined by

$$||A||_{B(X)} := \sup_{x \in X} \frac{||Ax||}{||x||}. \tag{A.1.6}$$

A family of bounded operators $\{U(a) : a \in \mathbb{R}^n\}$ is referred to as *strongly continuous* if for all $a \in \mathbb{R}^n$

$$\lim_{b \to a} ||U(a) - U(b)||_{B(X)} = 0 \tag{A.1.7}$$

and *weakly continuous* if

$$\lim_{b \to a} |\langle x, [U(a) - U(b)]y \rangle| = 0 \qquad \forall x, y \in X. \tag{A.1.8}$$

Note that by the Cauchy-Schwarz inequality on X, a strongly continuous family is always weakly continuous. If $\lim_{a \to \pm\infty} ||U(a) - A||_{B(X)} = 0$ for $A \in B(X)$ we say that $U(a)$ *converges strongly* to A and write

$$s - \lim_{a \to \pm\infty} U(a) = A \tag{A.1.9}$$

and call A its *strong limit*. If $\lim_{a \to \pm\infty} \langle x, [U(a) - A]y \rangle = 0$ for all $x, y \in X$ we say that $U(a)$ *converges weakly* to A and write

$$w - \lim_{a \to \pm\infty} U(a) = A \tag{A.1.10}$$

and call A its *weak limit*.

A confusing convention is to call a representation $U(a, \Lambda)$ of the Poincaré group $\mathscr{P}_+^\uparrow$ on a Hilbert space $\mathfrak{H}$ *strongly continuous*, if the corresponding family of operators is weakly continuous. This terminology stems from the theory of operator semigroups. In this context, the stronger notion is given by the condition

$$\lim_{(a,\Lambda) \to (0,\mathrm{id})} ||U(a, \Lambda) - \mathrm{id}_{\mathfrak{H}}||_{B(\mathfrak{H})} = 0, \tag{A.1.11}$$

with respect to the operator norm $|| \cdot ||_{B(\mathfrak{H})}$ and is referred to as *norm continuity*. However, the Poincaré group is generally not asked to be norm-continuous.

A subspace $N \subset X$ is called *invariant* if $U(a)N \subset N$ for all $a \in \mathbb{R}^n$. The family $\{U(a)\}$ is called *irreducible* if the only invariant subspaces are $\{0\}$ and X. We say that an operator A is densely defined, if its domain $D(A)$ is dense in X. A is called *symmetric* (or *hermitian*), if

$$\langle x, Ay \rangle = \langle Ax, y \rangle \qquad \forall x, y \in D(A). \tag{A.1.12}$$

If a symmetric operator is densely defined, ie $D(A) \subset X$ is dense, and B is a densely defined operator such that

$$\langle x, Ay \rangle = \langle Bx, y \rangle \qquad \forall x \in D(B), \forall y \in D(A) \tag{A.1.13}$$

then B is called the *adjoint operator* of A. We write $B = A^\dagger$. An operator A is referred to as *self-adjoint* if $D(A^\dagger) = D(A)$ and $A = A^\dagger$.

Closed operator. Note that if X is a Hilbert space with inner product $\langle\cdot,\cdot\rangle$, then $X \times X$ can be a given a Hilbert space structure by defining linear combinations through

$$\alpha(x_1,x_2)+\beta(y_1,y_2) := (\alpha x_1+\beta y_1, \alpha x_2+\beta y_2) \qquad \alpha,\beta\in\mathbb{C}. \tag{A.1.14}$$

The scalar product is given by $\langle x,y\rangle_{X^2} := \langle x_1,y_1\rangle + \langle x_2,y_2\rangle$ for $x,y\in X\times X$, explicitly written as $x=(x_1,x_2)$ and $y=(y_1,y_2)$. Let A be an operator on X and $D(A)\subset X$ its domain. The subspace

$$\Gamma(A) := \{\ (x,Ax)\mid x\in D(A)\ \}\subset X\times X \tag{A.1.15}$$

is called the *graph* of A. The operator A is called *closed* if $\Gamma(A)$ is complete with respect to the induced norm $||x||_{X^2} := \sqrt{\langle x,x\rangle_{X^2}}$ in $X\times X$. Let $\Gamma(A)$ be the graph of an operator A. This operator is called *closeable*, if there exists an operator $\overline{A}$ such that

$$\overline{\Gamma(A)} = \Gamma(\overline{A}), \tag{A.1.16}$$

ie the completion $\overline{\Gamma(A)} := \mathrm{CS}(\Gamma(A))/\sim$ of the pre-Hilbert space $\Gamma(A)$ is the graph of $\overline{A}$, which is said to be the *closure* of A. Finally, a symmetric operator A is called *essentially self-adjoint* if its closure $\overline{A}$ is self-adjoint.

A.2. Concise introduction to Hopf algebras

A Hopf algebra is a set equipped with unusually many algebraic structures. The reader is asked for patience during the course of the following passages.

Tensor space. Let A,B be vector spaces and $\{a_i : i\in I\}$ a basis in A, $\{b_j : j\in J\}$ a basis in B with index sets $I,J\subset\mathbb{N}$, not necessarily finite. The *tensor space* $A\otimes B$ is the vector space over $\mathbb{Q}$ spanned by pairs of the form $e_{jk}=a_j\otimes b_k$ with the following properties:

$$\begin{gathered}\lambda a_j\otimes b_k = a_j\otimes\lambda b_k \qquad \forall\lambda\in\mathbb{F},\\ a_j\otimes b_k + a_i\otimes b_k = (a_j+a_i)\otimes b_k, \qquad a_j\otimes b_i + a_j\otimes b_k = a_j\otimes(b_i+b_k).\end{gathered} \tag{A.2.1}$$

This implies $A\cong\mathbb{F}\otimes A\cong A\otimes\mathbb{F}$ since for example $\lambda\otimes a = 1\otimes\lambda a$ for any $a\in A$ and any $\lambda\in\mathbb{F}$, ie because the basis in $\mathbb{F}$ is simply given by 1. We will always write $\lambda\otimes a=\lambda a$ and identify such objects if they arise. For two linear maps $f\colon A\to A$ and $g\colon B\to B$ we can define a linear map $f\otimes g\colon A\otimes B\to A\otimes B$ by setting

$$(f\otimes g)(a\otimes b) := f(a)\otimes g(b). \tag{A.2.2}$$

If $A=\mathbb{F}$, then $(f\otimes g)(\lambda\otimes b) = f(\lambda)g(b) = \lambda f(1)g(b) = f(1)g(\lambda b) = g(\lambda f(1)b)$ by linearity of f and g.

Algebra. We define an *algebra* A as a vector space with an associative product, distributive with respect to the addition and containing a neutral element 1_A called unit. We view the product as a linear map $m\colon A\otimes A\to A$ and write the product of two elements $x,y\in A$ as

$$m(x\otimes y) = xy, \tag{A.2.3}$$

ie as a simple juxtaposition. For the product, linearity means[3]

$$m(x\otimes y+\lambda w\otimes z) = m(x\otimes y)+m(\lambda w\otimes z) = xy+\lambda wz \tag{A.2.4}$$

where $\lambda\in\mathbb{F}$ and $x,y,w,z\in A$. For the unit map, we have $1_A a = a1_A = a$. Associativity can be expressed in the form $m(m\otimes\mathrm{id}) = m(\mathrm{id}\otimes m)$ because of

$$m(m(x\otimes y)\otimes z) = m(xy\otimes z) = (xy)z = x(yz) = m(x\otimes yz) = m(x\otimes m(y\otimes z)). \tag{A.2.5}$$

[3]This property is normally not part of the definition of the product, but employed here for reasons to be understood later in this section.

The *tensor algebra* of two algebras A and B is the tensor space $A \otimes B$ with associative product

$$(a \otimes b)(a' \otimes b') = aa' \otimes bb'. \tag{A.2.6}$$

The reader may check that associativity of this product is inherited. We often need the so-called unit map $u\colon \mathbb{F} \to A$ which simply takes a scalar λ to $\lambda 1_A$.

Definition A.1 (Algebra). *An associative unital* $\mathbb{F}$-algebra *is a triple* (A, m, u) *consisting of a vector space A over $\mathbb{F}$,*

(1) *an associative linear map $m\colon A \otimes A \to A$ called product and*
(2) *a linear map $u\colon \mathbb{F} \to A$, $\lambda \mapsto u(\lambda) = \lambda 1_A$, refered to as unit map.*

In this work, all algebras are both unital and associative. Therefore, where we speak of an algebra, we assume these two properties without referring to them explicitly. Examples are polynomials $\mathbb{F}[X]$ in a variable X, continuous functions $\mathscr{C}^0(\mathbb{F})$ (here: $\mathbb{F} = \mathbb{R}, \mathbb{C}$) and matrices $\mathbb{F}^{n\times n}$. This should be familiar to the reader. The only perhaps new aspect is the linearity of the product.

Coalgebra. Given an algebra A, we may be interested in the dual vector space A^* of linear functionals $A \to \mathbb{F}$, also known as *covectors*. Let $f \in A^*$. We write $f(a) = \langle f, a\rangle$ for its action on a vector $a \in A$. What is the map dual to the product m? If we denote it by Δ, it has to satify

$$\langle f, m(a \otimes b)\rangle = \langle \Delta(f), a \otimes b\rangle, \tag{A.2.7}$$

and must surely map A^* to $A^* \otimes A^* \simeq (A \otimes A)^*$, where $\langle f \otimes g, a \otimes b\rangle := \langle f, a\rangle\langle g, b\rangle$. A quick calculation shows that associativity of the product requires

$$(\mathrm{id} \otimes \Delta)\Delta = (\Delta \otimes \mathrm{id})\Delta \tag{A.2.8}$$

to hold on A^*. The reader is encouraged to prove that this property follows from the associativity of the product m. It is known as *coassociativity*. The linear map Δ is called *coproduct*. The unit map $u\colon A \to \mathbb{F}$ does also have a dual which we denote by ε and refer to as *counit*. Because of

$$\langle f, 1_A\rangle = \langle f, u(1)\rangle = \langle \varepsilon(f), 1_A\rangle = \varepsilon(f)1 = \varepsilon(f) \tag{A.2.9}$$

it must map A^* to $\mathbb{F}$. Additionally, by

$$\langle f, a\rangle = \langle f, 1_A a\rangle = \langle f, u(1)a\rangle = \langle f, m(u(1) \otimes a)\rangle = \langle \Delta(f), u(1) \otimes a\rangle = \langle (\varepsilon \otimes \mathrm{id})\Delta(f), 1 \otimes a\rangle \tag{A.2.10}$$

where $1 \otimes a \cong a$ and the same for $a = a1_A = au(1)$ the counit $\varepsilon\colon A \to \mathbb{F}$ is required to fulfil

$$(\mathrm{id} \otimes \varepsilon)\Delta = (\varepsilon \otimes \mathrm{id})\Delta = \mathrm{id}\ . \tag{A.2.11}$$

In general, without having to be characterized as a dual, a vector space C equipped with a coproduct Δ and counit ε such that (A.2.11) is called *coalgebra*:

Definition A.2 (Coalgebra). *Let C be a vector space over $\mathbb{F}$. The triple (C, Δ, ε) is called coassociative counital* $\mathbb{F}$-coalgebra*, if it is equipped with*

(1) *a coassociative coproduct $\Delta\colon A \to A \otimes A$, a linear map with (A.2.11) and*
(2) *a counit $\varepsilon\colon A \to \mathbb{F}$, ie a linear map such that* $(\mathrm{id} \otimes \varepsilon)\Delta = (\varepsilon \otimes \mathrm{id})\Delta = \mathrm{id}$.

If not otherwise stated, a coalgebra is in this work tacitly assumed to be coassociative and counital. An example is the vector space $\mathbb{F}[X]$ of polynomials: the coproduct

$$\Delta(X^n) := \sum_{j=0}^{n} \binom{n}{j} X^j \otimes X^{n-j} \qquad (n \in \mathbb{N}_0) \tag{A.2.12}$$

which defines Δ uniquely. The counit is given by $\varepsilon(X^n) = 0$ for $n \geq 1$ and $\varepsilon(1) = 1$. It is a nice exercise to prove that these so-defined linear maps really do establish a coalgebra structure on

$\mathbb{F}[X]$ and also to find that the binomial coefficient in (A.2.12) can be dropped with no harm. Another example is the vector space $\mathbb{R}[\partial_x]$ of polynomials, where ∂_x is the usual differential operator acting on smooth functions $\mathbb{R} \to \mathbb{R}$. The structures Δ and ε are defined in the same way as for the variable X. The reader may try to prove the identity

$$\partial_x^n(f(x)g(x)) = \Delta(\partial_x^n)(f(x) \otimes g(x)). \tag{A.2.13}$$

for any smooth $f, g \in \mathscr{C}^\infty(\mathbb{R})$.

Bialgebra. We are now very close to a Hopf algebra. Consider again the algebra of polynomials $\mathbb{F}[X]$. We have seen that on the other hand, it can be equipped with a coalgebra structure. If we combine the structures of an algebra and that of a coalgebra, we have the ingredients of what is known as a *bialgebra* B if two conditions are fulfilled:

$$\Delta(ab) = \Delta(a)\Delta(b), \qquad \varepsilon(ab) = \varepsilon(a)\varepsilon(b), \tag{A.2.14}$$

in words: both coalgebra structures Δ and ε must respect the algebra structures, ie both are required to be multiplicative and linear. It is revealing to see that in the case $B = \mathbb{F}[X]$, the coproduct, as defined in (A.2.12), is not multiplicative if the binomial factor is omitted:

$$\Delta'(X^n) = \sum_{j=0}^{n} X^j \otimes X^{n-j} \neq \sum_{j=0}^{n} \binom{n}{j} X^j \otimes X^{n-j} = (1 \otimes X + X \otimes 1)^n = (\Delta(X))^n, \tag{A.2.15}$$

ie although the coproduct Δ' is perfectly fine for a coalgebra, it is not for a bialgebra.

Definition A.3 (Bialgebra). *A* bialgebra *is a quintuple* $(B, m, u, \Delta, \varepsilon)$ *in which* (B, m, u) *is an algebra,* (B, Δ, ε) *a coalgebra and the coalgebra's maps* Δ, ε *are multiplicative as illustrated by (A.2.14).*

Hopf algebra. Suppose H is a bialgebra, ie given by the quadruple $(H, m, u, \Delta, \varepsilon)$. With these structures, we can now establish an associative bilinear operation on the space $\mathcal{L}(H)$ of linear maps from H to itself by setting

$$f \star g := m(f \otimes g)\Delta, \tag{A.2.16}$$

which means $(f \star g)(x) = m(f \otimes g)\Delta(x) = m(f \otimes g)(\sum_{(x)} x' \otimes x'') = \sum_{(x)} f(x')g(x'') \in H$ if we use a variant of Sweedler's notation for the coproduct given by $\Delta(x) = \sum_{(x)} x' \otimes x''$. This operation is called *convolution product.* Note that $f \star g \colon H \to H$ is again linear and the composition $e = u \circ \varepsilon$ turns out to be the neutral element of the convolution:

$$(f \star e)(x) = \sum_{(x)} f(x') \underbrace{e(x'')}_{u(\varepsilon(x''))} = \sum_{(x)} f(x')\varepsilon(x'')1_H = \sum_{(x)} f(x'\varepsilon(x'')) = f(\sum_{(x)} x'\varepsilon(x'')) = f(x) \tag{A.2.17}$$

where $\sum_{(x)} x'\varepsilon(x'') = (\mathrm{id} \otimes \varepsilon)\Delta(x) = \mathrm{id}(x)$ is a property of the coalgebra. $(e \star f)(x) = f(x)$ goes along the same lines. We now ask whether there is an inverse of a map $f \in \mathcal{L}(H)$ with respect to the convolution product. In particular, whether it exists for the identity map $f = \mathrm{id}$. If the answer to this latter question is yes, we call it the *antipode* (or coinverse) S and write its defining property as

$$S \star \mathrm{id} = \mathrm{id} \star S = \mathrm{e}. \tag{A.2.18}$$

Now, there we are. A bialgebra H that has the luxury of an antipode is called *Hopf algebra*:

Definition A.4 (Hopf algebra). *A hextuple* $(H, m, u, \Delta, \varepsilon, S)$ *is called* Hopf algebra, *if the quintuple* $(H, m, u, \Delta, \varepsilon)$ *is a bialgebra and* $S \colon H \to H$ *an antipode, ie a linear map such that* $S \star \mathrm{id} = \mathrm{id} \star S = \mathrm{e}$.

We shall see that in many cases, the antipode can be defined recursively. Take again the example $H = \mathbb{F}[X]$ where

$$S(X^n) = -X^n - \sum_{j=1}^{n-1} \binom{n}{j} S(X^j)X^{n-j} = -X^n - \sum_{j=1}^{n-1} \binom{n}{j} X^j S(X^{n-j}) \tag{A.2.19}$$

for a monomial X^n with $n \geq 1$ which follows from $(\mathrm{id} \star S)(X^n) = (S \star \mathrm{id})(X^n) = e(X^n) = 0$, as can be seen here:
(A.2.20)

$$0 = (\mathrm{id}\star S)(X^n) = \sum_{j=0}^{n} \binom{n}{j} \mathrm{id}(X^j)S(X^{n-j}) = \mathrm{id}(1)S(X) + \sum_{j=1}^{n-1} \binom{n}{j} \mathrm{id}(X^j)S(X^{n-j}) + \mathrm{id}(X)S(1),$$

and because $1 = e(1) = (\mathrm{id}\star S)(1) = \mathrm{id}(1)S(1) = S(1)$, on account of $\Delta(1) = 1 \otimes 1$ which follows from (A.2.12), the antipode preserves the unit $S(1) = 1$.

A.3. Convolution algebra and group

Let $\mathcal{L}(C, A)$ be the set of linear maps from a coalgebra $(C, \Delta_C, \varepsilon_C)$ to an algebra (A, m_A, u_A). By virtue of the structures on both spaces, the *convolution* of two linear maps $f, g \in \mathcal{L}(C, A)$, given by

$$f \star g := m_A(f \otimes g)\Delta_C \ , \tag{A.3.1}$$

is an associative bilinear operation on $\mathcal{L}(C, A)$. The map $\mathrm{e} := u_A \circ \varepsilon_C$ is the neutral element with respect to $\star$. This makes the linear space $\mathcal{L}(C, A)$ into an algebra:

Proposition A.5 (Convolution algebra). *$\mathcal{L}(C, A)$ is an algebra with respect to $\star$, the* convolution algebra.

Proof. We take $f, g, h \in \mathcal{L}(C, A)$ and first check associativity,

$$\begin{aligned}(f \star g) \star h &= m_A((f \star g) \otimes h)\Delta_C = m_A(m_A(f \otimes g)\Delta_C \otimes h)\Delta_C \\ &= m_A(m_A \otimes \mathrm{id})((f \otimes g) \otimes h)(\Delta_C \otimes \mathrm{id})\Delta_C \\ &= m_A(\mathrm{id} \otimes m_A)(f \otimes (g \otimes h))(\mathrm{id} \otimes \Delta_C)\Delta_C \\ &= m_A((f \otimes m_A(g \otimes h)\Delta_C)\Delta_C = m_A((f \otimes (g \star h))\Delta = f \star (g \star h).\end{aligned} \tag{A.3.2}$$

Furthermore, the property $(\mathrm{id}_C \otimes \varepsilon_C)\Delta_C = \mathrm{id}_C = (\varepsilon_C \otimes \mathrm{id}_C)\Delta_C$ can be used to show that the linear map $\mathrm{e}\colon C \to A$ is the neutral element: take $x \in C$ and compute

$$(f \star \mathrm{e})(x) = m_A(f \otimes u_A)(\mathrm{id}_C \otimes \varepsilon_C)\Delta_C(x) = m_A(f \otimes u_A)(x \otimes 1_{\mathbb{F}}) = f(x) \tag{A.3.3}$$

and likewise for $(\mathrm{e} \star f)(x) = f(x)$. □

Naturally, one can define $\star$-powers by setting $f^{\star 0} := \mathrm{e}$, $f^{\star 1} := f$ and $f^{\star n+1} := f \star f^{\star n}$ recursively. For $f \in \mathcal{L}(C, A)$ there may be a map $h \in \mathcal{L}(C, A)$ such that $f \star h = h \star f = \mathrm{e}$, ie a $\star$-inverse of f. This is guaranteed if we replace the coalgebra C by a graded connected bialgebra $(B, m_B, u_B, \Delta_B, \varepsilon_B)$ and f satisfies $f(1_B) = 1_A$. The grading is a decomposition into subspaces $B = \bigoplus_{n\geq 0} B_n$ such that

$$\Delta_B(B_n) \subset \bigoplus_{j=0}^{n} B_j \otimes B_{n-j} \qquad \text{(grading of a coalgebra)} \tag{A.3.4}$$

and

$$m_B(B_n \otimes B_m) \subset B_{n+m} \qquad \text{(grading of an algebra).} \tag{A.3.5}$$

'Connected' (grading) means that $B_0 = \mathbb{F}1_B$. The connected grading guarantees that for every element $x \in B$ there exists an $N > 0$ such that

$$(\mathrm{e} - f)^{\star n}(x) = 0 \qquad \forall n > N, \tag{A.3.6}$$

because if one applies the coproduct often enough, and it does not matter which part of a tensor product it is acted on (by coassociativity), there will be a 1_B in every term and then $(\mathrm{e} - f)(1_B) = \mathrm{e}(1_B) - f(1_B) = 0$ ensures that the von Neuman series terminates, hence exists.

Proposition A.6 (Convolution group). *The subset*

$$\mathcal{L}(B, A)^\times := \{f \in \mathcal{L}(B, A) | f(1_B) = 1_A\} \subset \mathcal{L}(B, A) \tag{A.3.7}$$

is a group *with respect to $\star$, called the* convolution group[4]*, ie for every map $f \in \mathcal{L}(B, A)^\times$ there exist a linear map $f^{\star -1}$ such that*

$$f \star f^{\star -1} = f^{\star -1} \star f = \mathrm{e} \tag{A.3.8}$$

and $f^{\star -1}(1_B) = 1_A$, given by the von Neumann series $f^{\star -1} = \sum_{n \geq 0}(\mathrm{e} - f)^{\star n}$.

Proof. Take $x \in B$. Then there is an $N > 0$ such that $(e - f)^{\star n} = 0$ for all $n > N$. Then, using the shorthand $\Delta(x) = x' \otimes x''$, we compute

$$\begin{aligned}
(f^{\star -1} \star f)(x) &= \sum_{n\geq 0}(\mathrm{e} - f)^{\star n}(x')f(x'') = \sum_{n\geq 0}(\mathrm{e} - f)^{\star n}(x')(\mathrm{e}(x'') - [\mathrm{e}(x'') - f(x'')]) \\
&= \sum_{n\geq 0}(\mathrm{e} - f)^{\star n}(x')\mathrm{e}(x'') - \sum_{n\geq 0}(\mathrm{e} - f)^{\star n}(x')[\mathrm{e}(x'') - f(x'')] \\
&= \sum_{n\geq 0}(\mathrm{e} - f)^{\star n}(x) - \sum_{n\geq 0}(\mathrm{e} - f)^{\star n+1}(x) = (\mathrm{e} - f)^{\star 0}(x) = \mathrm{e}(x)\ .
\end{aligned} \tag{A.3.9}$$

This works equally well with $(f \star f^{\star -1})(x)$. □

A.4. Algebraic Birkhoff decomposition and convolution group

Let $f \in \mathcal{L}(B, A)^\times$ and $A = A_- \oplus A_+$ be a decomposition into linear subspaces. A pair of maps $f_\pm \in \mathcal{L}(B, A)^\times$ is called *algebraic Birkhoff decomposition* of f with respect to the decomposition $A_\pm$ if

$$f_\pm(\ker \varepsilon_B) \subset A_\pm \qquad \text{and} \qquad f = f_-^{*-1} * f_+\ . \tag{A.4.1}$$

We denote the projector onto $\ker \varepsilon_B$ by P_B. Given two subspaces, the Birkhoff decomposition always exists and is unique.

Theorem A.7 (Birkhoff decomposition). *Let $f \in \mathcal{L}(B, A)^\times$ and $A = A_- \oplus A_+$ be a decomposition into subspaces with projector $R : A \to A_-$. Then, the Birkhoff decomposition $f_\pm \in \mathcal{L}(B, A)^\times$ is uniquely defined by the recursive relations*

$$f_-(x) = -R[(f_- * fP_B)(x)], \tag{A.4.2}$$

for every $x \in \ker \varepsilon_B$ and $f_+ := f_- * f$.

Proof. First existence. We define the linear map by setting $f_-(1_B) := 1_A$ and using (A.4.2) which determines f_- uniquely due to

$$(f_- * fP_B)(x) \in f(x) + m_A(f_- \otimes f)(\textstyle\bigoplus_{j=1}^{n-1} B_j \otimes B_{n-j})\ . \tag{A.4.3}$$

[4] In this view, ignoring the linear structure, the convolution algebra is a monoid.

$f_-(\ker\varepsilon_B) \subset A_-$ is satisfied by definition (due to the projector R). $f_+ \in \mathcal{L}(B,A)^\times$ because of $f_+(1_B) = f_-(1_B)f(1_B) = 1_A$. On account of

$$\begin{aligned} f_+(x) &= (f_- * f)(x) = (f_- * fP_B)(x) + f_-(x) = (f_- * fP_B)(x) - R[(f_- * fP_B)(x)] \\ &= [\mathrm{id}_B - R](f_- * fP_B)(x) \in A_+ \end{aligned} \tag{A.4.4}$$

for $x \in \ker\varepsilon_B$ we have $f_+(\ker\varepsilon_B) \subset A_+$, because $[\mathrm{id}_B - R]$ projects onto A_+. Now Uniqueness: any Birkhoff decomposition $f_\pm$ satisfies (A.4.2), that is, if we take any $x \in \ker\varepsilon_B$, then

$$-R[(f_- * fP_B)(x)] = -R[(f_- * f)(x) - f_-(x)] = -R[f_+(x) - f_-(x)] = R[f_-(x)] = f_-(x)\,. \tag{A.4.5}$$

Because this recursive relation determines a map f_- uniquely, the Birkhoff decomposition is unique. ☐

A concrete example is the algebra $\mathbb{C}[z^{-1}, z]]$ of Laurent series without essential singularities, where a decomposition is given by

$$A_- = z^{-1}\mathbb{C}[z^{-1}]\,, \qquad A_+ = \mathbb{C}[[z]]\,. \tag{A.4.6}$$

The set A_- consists of all polynomials in the variable z^{-1} having no constant part, which implies $\lim_{z\to\infty} f(z) = 0$ for all $f \in A_-$. Note that both subspaces are subalgebras but A_- is not unital, whereas A_+ is.

A.5. Character group

If we replace the connected bialgebra B by a connected Hopf algebra H, the convolution group has a subset

$$\mathrm{Ch}(H,A) := \{f \in \mathcal{L}(H,A)^\times \mid f(xy) = f(x)f(y)\ \forall x,y \in H\} \tag{A.5.1}$$

of multiplicative maps in which the inverse $f^{\star-1}$ of an element $f \in \mathrm{Ch}(H,A)$ is given by the linear map $fS := f \circ S$, where S is the Hopf algebra's antipode:

$$\begin{aligned} (fS \star f)(x) &= f(S(x'))f(x'') = f(S(x')x'') = f(\mathrm{e}(x)) = f(u_H\varepsilon_H(x)) = f(\varepsilon_H(x)1_H) \\ &= \varepsilon_H(x)1_A = u_A\varepsilon_H(x), \end{aligned} \tag{A.5.2}$$

where $u_A\varepsilon_H := u_A \circ \varepsilon_H$ is the neutral element with respect to $\star$. Note that fS is not necessarily in the subset $\mathrm{Ch}(H,A)$! This shows the following calculation:

$$fS(xy) = f(S(xy)) = f(S(y)S(x)) = fS(y)\ fS(x) \tag{A.5.3}$$

which may not be equal to $fS(x)fS(y)$. However, if the target algebra A is commutative, it is:

Proposition A.8 (Character group). *Let A be commutative. Then,* $\mathrm{Ch}(H,A) \subset \mathcal{L}(H,A)^\times$ *is a subgroup of the convolution group.*

Proof. We have seen that $f^{\star-1} = fS \in \mathrm{Ch}(H,A)$. For $f, g \in \mathrm{Ch}(H,A)$ we find

$$\begin{aligned} (f \star g)(xy) &= f(x'y')g(x''y'') = f(x')f(y')g(x'')g(y'') = f(x')g(x'')f(y')g(y'') \\ &= (f \star g)(x)(f \star g)(y)\,, \end{aligned} \tag{A.5.4}$$

that is, $f \star g \in \mathrm{Ch}(H,A)$, where we have used the shorthand notation $\Delta(x) = x' \otimes x''$. ☐

This subgroup is named *character group*. Its elements are called *Hopf algebra characters* or just *Hopf characters*. In view of the Birkhoff decomposition of elements in the convolution group, we may ask what the state of affair is for characters: does the Birkhoff decomposition lie entirely in $\mathrm{Ch}(H,A)$? The next proposition tells us that for this to be true, the projector $R\colon A \to A_-$ must be Rota-Baxter.

Proposition A.9. *In the setup of Theorem A.7, let f be a Hopf character and the projector R be a Rota-Baxter operator, ie such that*

$$R[ab] + R[a]R[b] = R[aR[b] + R[a]b] \tag{A.5.5}$$

for all $a, b \in A$, where A is commutative. This is garanteed if $A_\pm$ are subalgebras[5]*. Then the Birkhoff decomposition $f_\pm$ of f consists of Hopf characters.*

Proof. To understand the last assertion, ie that a projector onto subalgebras is always Rota-Baxter, one can easily check that (A.5.5) is fulfilled in the possible cases $a \in \ker R = A_+$, $b \in \operatorname{im} R = A_-$, vice versa and so on. The proof is inductive with respect to the grading of H. For $H_0 = \mathbb{Q}1_H$. Assume $f_\pm$ are multiplicative on $\bigoplus_{j=0}^n H_j$. Then, choose $x, y \in H$ such that $xy \in H_{n+1}$. We use the abbreviation

$$\overline{f} := f_- \star fP \tag{A.5.6}$$

in the following computation. $P\colon H \to \bigoplus_{j\geq 1} H_j$ is a projector. Then,

(A.5.7)

$$\begin{aligned}
&f_-(xy) = -R[(f_-(x'y')fP(x''y'')] \overset{(*)}{=} -R[f_-(x')f_-(y')fP(x''y'')]\\
&= -R[f_-(x')f_-(y')f(x''y'') - f_-(x)f_-(y)] = -R[f_-(x')f_-(y')f(x'')f(y'') - f_-(x)f_-(y)]\\
&= -R[f_-(x')f(x'')f_-(y')f(y'') - f_-(x)f_-(y)] = -R[(f_- \star f)(x)(f_- \star f)(y) - f_-(x)f_-(y)]\\
&= -R[(\overline{f}(x) + f_-(x))(\overline{f}(y) + f_-(y)) - f_-(x)f_-(y)]\\
&= -R[\overline{f}(x)\overline{f}(y) + \overline{f}(x)f_-(y) + f_-(x)\overline{f}(y)] = -R[\overline{f}(x)\overline{f}(y) - \overline{f}(x)R\overline{f}(y) - R\overline{f}(x)\overline{f}(y)]\\
&= R[\overline{f}(x)]R[\overline{f}(y)] = f_-(x)f_-(y),
\end{aligned}$$

where we have used in $(*)$ that $x'y' \in H_{n+1}$ only if $x' = x, y' = y$, that is, only if $x''y'' = 1_H$, which does not appear in the sum due to the presence of the projector P. Hence f_- is multiplicative in that step of the calculation. Then so is $f_+ = f_- \star f$ (by Proposition A.8). □

A.6. Ideals

Ideals. Let A be an algebra over a field $\mathbb{F}$. A subspace $I \subset A$ is called *left ideal* if $AI \subset I$, and *right ideal* $IA \subset I$, ie if $ax \in I$ for a left and $xa \in I$ for a right ideal whenever $x \in I$ and $a \in A$. If both conditions are satisfied, then I is called (two-sided) *ideal*. Note that, trivially, (left/right) ideals are subalgebras by definition and, of course, if the product is commutative, both right and left ideals coincide.

Here is an example. Take the set of polynomials $A = \mathbb{F}[X]$ in one variable. Let $c \in \mathbb{F}$ be any number. The set of polynomials defined by

$$I_c := \{\ p \in \mathbb{F}[X] \mid p(c) = 0\ \} \tag{A.6.1}$$

clearly form a subspace and, surely, a subalgebra. It is moreover an ideal, since $q(c)p(c) = 0$ for $p \in I_c$, even if $q(c) \neq 0$ for $q \notin I_c$. We can in fact choose any polynomial $q \in A$ and generate an ideal

$$(q) := \{aq \mid a \in A\}, \tag{A.6.2}$$

known as *principle ideal*. This really is an ideal since any $r \in (q)$ is of the form $r = aq$ and we can multiply it with anything $w \in A$ and find $wr = waq \in (q)$ since $wa \in A$.

[5]Not necessarily unital algebras!

Hopf ideals. A less trivial question is whether an ideal $I \subset H$ is also a so-called *coideal* of a Hopf algebra H, ie if

$$\Delta(I) \subset I \otimes H + H \otimes I. \tag{A.6.3}$$

Furthermore, we may ask whether the antipode respects it: $S(I) \subset I$. If these two conditions are satisfied, I is referred to as *Hopf ideal.* Let us see if $I_c \subset H = \mathbb{F}[X]$ defined in (A.6.1) is a coideal and maybe even Hopf. Take $p(X) = X - c \in I_c$. The coproduct gives

$$\Delta(p(X)) = X \otimes 1 + 1 \otimes X - c1 \otimes 1 = (X - c) \otimes 1 + 1 \otimes X. \tag{A.6.4}$$

Only for $c = 0$ is this an element in $H \otimes I_0 + I_0 \otimes H$. We choose $c = 0$. The coproduct of a monomial $X^n \in I_0$ for $n \neq 0$ is

$$\Delta(X^n) = \sum_{j=0}^{n} \binom{n}{k} X^j \otimes X^{n-j} = 1 \otimes X + X \otimes 1 + \sum_{j=1}^{n-1} \binom{n}{k} X^j \otimes X^{n-j} \tag{A.6.5}$$

where $1 \otimes X + X \otimes 1 \in H \otimes I + I \otimes H$ and the remainder is actually in $I \otimes I \subset I \otimes H + H \otimes I$. Since this holds for all monomials, we have for any polynomial $p \in I_0$

$$\Delta(p(X)) \subset H \otimes I_0 + I_0 \otimes H \tag{A.6.6}$$

since $p(X)$ must be a linear combination of monomials X^n with $n \neq 0$. Therefore $I_0 \subset H = \mathbb{F}[X]$ is indeed a coideal. One can show that $S(I_0) \subset I_0$ by the antipode's multiplicativity:

$$S(X^n) = S(X)^n = (-X)^n = (-1)^n X^n \in I_0, \tag{A.6.7}$$

where $S(X) = -X$ follows from $0 = (S * \mathrm{id})(X) = m(S \otimes \mathrm{id})\Delta(X) = S(X)1 + S(1)X$ and $S(1) = 1$. We conclude: the ideal I_c is a Hopf ideal iff $c = 0$.

A.7. Graded and differential algebras

Let in the following A be an algebra over $\mathbb{R}$ and $A = \bigoplus_{j \geq 0} A_j$ a direct sum of subspaces.

DEFINITION A.10 (Graded algebra). *The direct sum $A = \bigoplus_{j \geq 0} A_j$ is called* grading *of A, if*

$$A_i A_j \subset A_{i+j} \tag{A.7.1}$$

for all indices. The algebra A is then a graded algebra*. The elements in the subspaces of the grading are called* homogeneous*, in particular, $x \in A_j$ is said to be* homogeneous of degree j.

An example is the polynomial algebra $\mathbb{R}[X]$ with grading $\mathbb{R}[X] = \bigoplus_{n \geq 0} \mathbb{R}X^n$, eg $f(X) = 3X^2$ is homogeneous of degree 2, whereas $g(X) = 2X + 4X^5$ is not, yet its components with respect to the grading are: $g_1(X) = 2X$ and $g_2(X) = 4X^5$ are homogeneous of degree 1 and 5, respectively.

The index set of a grading can easily be generalised: it needs only be a monoid, ie a set with a binary operation and a neutral element. In particular, one can consider a grading with index set $\mathbb{Z}^n$. This is needed if more than one polynomial variable is used. For example $\mathbb{R}[X, Y]$ needs $\mathbb{N}_0^2$ as index set.

DEFINITION A.11 (Differential algebra). *A linear map $D\colon A \to A$ is called* derivation *if*

$$D(ab) = D(a)b + aD(b) \tag{A.7.2}$$

for all $a, b \in A$. An algebra equipped with a derivation is referred to as differential algebra.

A familiar example are smooth functions in one variable x with $D = \partial_x$ as a derivation for which (A.7.2) then is just what is known as the *Leibniz rule* or *product rule*.

A.8. Basics of Borel summation

An asymptotic series is a Taylor series expansion around, say, the origin in $\mathbb{C}$ which may or may not converge. The information content that it still has, however, is given by the values of the derivatives of a smooth function at the origin.

Definition A.12 (Gevry-n). *An asymptotic series* $\sum_{k\geq 0} a_k x^k$ *is called* Gevry-n*, if*

$$\sum_{k\geq 0} \frac{a_k}{(k!)^n} x^k \tag{A.8.1}$$

has nonzero radius of convergence.

Definition A.13 (Borel transform). *Let* $f(x) = \sum_{k\geq 0} a_k x^k$ *be Gevry-*1*. The series*

$$\mathscr{B}f(\zeta) := \sum_{k\geq 0} \frac{a_k}{k!} \zeta^k \tag{A.8.2}$$

is called the Borel transform *of* $f(x)$*. By* $\mathtt{cont}\mathscr{B}[f]$ *we denote its meromorphic continuation along the positive real axis* $\mathbb{R}_+ = (0, \infty)$*.*

Definition A.14 (Borel-Laplace transform). *Let* $h : \mathbb{R}_+ \to \mathbb{R}$ *be continuous and of at most exponential growth, ie there exist constants* $C, a > 0$ *such that* $|h(x)| \leq Ce^{ax}$ *for all* $x \geq 0$*. The integral*

$$\mathscr{L}[h](x) := \frac{1}{x} \int_0^\infty d\zeta \; e^{-\zeta/x} h(\zeta) \tag{A.8.3}$$

is called Borel-Laplace transform *of* h*.*

Definition A.15 (Borel summability). *An asymptotic series* $f(x) = \sum_{k\geq 0} a_k x^k$ *of class Gevry-1 is called* Borel-summable *if the Borel-Laplace transform of its Borel transform, given by*

$$\mathscr{L}[\mathscr{B}f](z) := \frac{1}{z} \int_0^\infty d\zeta \; e^{-\zeta/z} \mathtt{cont}\mathscr{B}[f](\zeta) \tag{A.8.4}$$

exists for some $z \neq 0$*. The function* $\mathscr{L}[\mathscr{B}f](z)$ *is also referred to as the* Borel sum *of* f*.*

Note that Laplace transforms converge in right half planes. These definitions are motivitated by the following (in parts) formal computation

$$\sum_{n\geq 0} a_n x^n = \sum_{n\geq 0} \frac{a_n}{n!} \frac{1}{x} \underbrace{\int_0^\infty d\zeta \; \zeta^n e^{-\zeta/x}}_{=n!x^{n+1}} = \frac{1}{x} \int_0^\infty d\zeta \; \zeta^n e^{-\zeta/x} \sum_{n\geq 0} \frac{a_n}{n!} \zeta^n. \tag{A.8.5}$$

An standard textbook example is given by the series $f(x) = \sum_{n\geq 0} (-1)^n n! x^n$. Its Borel transform is

$$\mathscr{B}f(\zeta) = \sum_{n\geq 0} (-1)^n \zeta^n \qquad \Longrightarrow \qquad \mathtt{cont}\mathscr{B}f(\zeta) = \frac{1}{1+\zeta}. \tag{A.8.6}$$

This function has a Borel-Laplace transform given by

$$\mathscr{L}[\mathscr{B}f](x) = \frac{1}{x} \int_0^\infty \frac{e^{-\zeta/x}}{1+\zeta} = \frac{1}{x} \int_{1/x}^\infty \zeta^{-1} e^{-\zeta} = \frac{1}{x} \Gamma(0; 1/x), \tag{A.8.7}$$

where $\Gamma(z; s) := \int_s^\infty d\zeta \; \zeta^{z-1} e^{-\zeta}$ for $z, s \geq 0$ is the well-known *upper incomplete gamma function.*

Finally, we cite Watson's theorem, one of the first results on Borel summability (see [**Sok79**] and references there).

THEOREM A.16 (Watson). *Let $f(z)$ be analytic in a sector*

$$S_\varepsilon(R) = \{z \in \mathbb{C} : 0 < |z| < R, |\arg(z)| < \frac{\pi}{2} + \varepsilon\} \tag{A.8.8}$$

for some $\varepsilon > 0$ and let

$$f(z) = \sum_{k=0}^{N-1} a_k z^k + R_N(z), \tag{A.8.9}$$

with $|R_N(z)| < A\sigma^N N!|z|^N$ uniformly in N and $z \in S_\varepsilon(R)$ for fixed $\sigma, A > 0$. Then

- *$\mathscr{B}f(\zeta) = \sum_{k\geq 0} \frac{a_k}{k!}\zeta^k$ converges in a circle $|\zeta| < 1/\sigma$.*
- *$\mathscr{B}f$ has an analytic continuation $\texttt{cont}_{S_\varepsilon}\mathscr{B}f$ to the sector $S_\varepsilon = \{\arg(\zeta) < \varepsilon\}$ and*
- *the Borel sum $\mathscr{L}[\mathscr{B}f](z)$ is absolutely convergent in the circle $C_R = \{\Re(z) > 1/R\}$ and*

$$f(z) = \mathscr{L}[\mathscr{B}f](z) \qquad z \in C_R. \tag{A.8.10}$$

Notice how that the flat function $f(z) = e^{-1/z}$ fails to fulfil the hypothesis (A.8.9): there is a path along which one can approach zero such that the function does not approach zero: we take the curve $\{x + iy : y = \sqrt{x}, x > 0\}$ and watch what happens as we let $x \downarrow 0$:

$$f(x + i\sqrt{x}) = \exp\left(-\frac{1}{1+x}\right)\exp\left(\frac{i}{x^{3/2} + x^{1/2}}\right). \tag{A.8.11}$$

This does not approach zero but oscillates rampantly.

APPENDIX B

Miscellaneous

This appendix chapter contains some proofs of assertions used in the main text. We have decided to relegate them to this place because of their technicalness.

B.1. Baumann's theorem

THEOREM B.1 (Baumann). *Let $n \geq 4$ be the space dimension and $\varphi(t,\cdot)$ a scalar field with conjugate momentum field $\pi(t,\cdot) = \partial_t\varphi(t,\cdot)$ such that the CCR (1.6.2) are obeyed and assume furthermore that $\dot\pi(t,\cdot) := \partial_t\pi(t,\cdot)$ exists. Then, if $\varphi(t,\cdot)$ has a vanishing vacuum expectation value and the provisions listed in the introduction of* [**Bau87**] *are satified, one has*

$$\dot\pi(t,f) - \varphi(t,\Delta f) + m^2\varphi(t,f) = 0 \tag{B.1.1}$$

for all $f \in \mathscr{D}(\mathbb{R}^n)$ and a parameter $m^2 > 0$.

PROOF. We only provide a sketch. To see the technical details, the reader is referred to Baumann's paper [**Bau87**]. First, take any $\Psi \in \mathfrak{D}$ from the dense domain of definition. Then

$$[\varphi(t,f), \dot\pi(t,g)]\Psi = 0 \tag{B.1.2}$$

is an easily obtained consequence of $\partial_t[\varphi(t,f),\pi(t,g)] = i\partial_t(f,g) = 0$ and causality of the momentum field, ie $[\pi(t,f),\pi(t,g)] = 0$ for any test functions $f,g \in \mathscr{D}(\mathbb{R}^n)$. The Jacobi identity then entails

$$[\varphi(t,f), [\pi(t,h), \dot\pi(t,g)]]\Psi = 0, \tag{B.1.3}$$

which is fairly straightforward. The hard part is to prove that $[\pi(t,f),[\pi(t,h),\dot\pi(t,g)]]\Psi = 0$. Baumann uses a partition of unity whose constituting test functions have compact support in ϵ-neighbourhoods that cover the support of the test functions f, g and h. Then he essentially shows that this double commutator vanishes with a power law for $\epsilon \downarrow 0$ that depends on the space dimension $n > 0$. If $n \geq 4$, then it vanishes. By virtue of the irreducibility of the field algebra $\{\varphi(t,\cdot),\pi(t,\cdot)\}$, one concludes from $[\varphi(t,f),[\pi(t,h),\dot\pi(t,g)]]\Psi = 0$ and $[\pi(t,f),[\pi(t,h),\dot\pi(t,g)]]\Psi = 0$ that

$$[\pi(t,h), \dot\pi(t,g)] = \langle\Omega|[\pi(t,h),\dot\pi(t,g)]\Omega\rangle, \tag{B.1.4}$$

ie that $[\pi(t,h),\dot\pi(t,g)]$ is a c-number. The next step is to write the rhs of (B.1.4) in terms of the Källen-Lehmann representation, ie a positive superposition of free commutator functions. Recall that the commutator function of the free field φ_0 with mass $m > 0$ is given by

$$[\varphi_0(t,\mathbf{x}),\varphi_0(s,\mathbf{y})] = \int \frac{d^4q}{(2\pi)^3}\,\delta_+(q^2-m^2)[e^{-iq\cdot(x-y)} - e^{+iq\cdot(x-y)}] =: D(t-s,\mathbf{x}-\mathbf{y};m^2), \tag{B.1.5}$$

where $\delta_+(q^2-m^2) := \theta(q_0)\delta(q^2-m^2)$ makes sure the particle is real and on-shell. One finds the corresponding representation of the commutator (B.1.4) by first differentiating the Källen-Lehmann representation

$$\langle\Omega|[\varphi(t,\mathbf{x}),\varphi(s,\mathbf{y})]\Omega\rangle = \int d\mu^2\ \rho(\mu^2)D(t-s,\mathbf{x}-\mathbf{y};\mu^2). \tag{B.1.6}$$

twice with respect to s which gives

$$\langle\Omega|[\varphi(t,\mathbf{x}),\dot{\pi}(s,\mathbf{y})]\Omega\rangle = \int d\mu^2\ \rho(\mu^2)[\Delta-\mu^2]D(t-s,\mathbf{x}-\mathbf{y};\mu^2) \tag{B.1.7}$$

because of $(\partial_s^2-\Delta+\mu^2)D(t-s,\mathbf{x}-\mathbf{y};\mu^2)=0$. Differentiating with respect to t and letting $t\to s$ yields the distribution

$$\langle\Omega|[\pi(t,\mathbf{x}),\dot{\pi}(t,\mathbf{y})]\Omega\rangle = -i\int d\mu^2\ \rho(\mu^2)[\Delta-\mu^2]\delta(\mathbf{x}-\mathbf{y}). \tag{B.1.8}$$

Applying this to two test functions $f,g\in\mathscr{D}(\mathbb{R}^n)$ gives

$$\langle\Omega|[\pi(t,f),\dot{\pi}(t,g)]\Omega\rangle = -i\int d\mu^2\ \rho(\mu^2)(f,[\Delta-\mu^2]g) = -i[(f,\Delta g)-m^2(f,g)] \tag{B.1.9}$$

with $m^2=\int d\mu^2\ \rho(\mu^2)\mu^2$ and $\int d\mu^2\ \rho(\mu^2)=1$. The latter is implied by the CCR, obtained from (B.1.6). Note that $m^2<\infty$ is warranted by the existence of the state $\dot{\pi}(g)\Omega$. Using the CCR, we can write the rhs of (B.1.9) as the commutator

$$-i[(f,\Delta g)-m^2(f,g)] = [\pi(t,f),\varphi(t,\Delta g)-m^2\varphi(t,g)] \tag{B.1.10}$$

and arrive at $[\pi(t,f),\dot{\pi}(t,g)-\varphi(t,\Delta g)+m^2\varphi(t,g)]=0$ because the commutator in the vacuum expectation value of the lhs of (B.1.9) is a c-number. The CCR in combination with (B.1.2) entail

$$[\varphi(t,f),\dot{\pi}(t,g)-\varphi(t,\Delta g)+m^2\varphi(t,g)]=0. \tag{B.1.11}$$

By the irreducibility of the field algebra, this means that $C:=\dot{\pi}(t,g)-\varphi(t,\Delta g)+m^2\varphi(t,g)$ is a c-number. The normalisation $\langle\Omega|\varphi(t,f)\Omega\rangle=0$ then yields $C=0$. □

B.2. Wightman's reconstruction theorem

We denote the algebra of Schwartz functions on Minkowski space $\mathbb{M}$ by $\mathscr{S}(\mathbb{M})$.

THEOREM B.2 (Reconstruction theorem). *Let $\{W_n\colon\mathscr{S}(\mathbb{M})^n\to\mathbb{C}\}$ be a family of tempered distributions satisfying the following set of properties.*

(1) POINCARÉ INVARIANCE. *$W_n(f_1,...,f_n)=W_n(\{a,\Lambda\}f_1,...,\{a,\Lambda\}f_n)$ for all Poincaré transformations $(a,\Lambda)\in\mathscr{P}_+^\uparrow$, where $(\{a,\Lambda\}f)(x):=f(\Lambda^{-1}(x-a))$.*

(2) SPECTRAL CONDITION. *W_n vanishes if one test function's Fourier transform has its support outside the forward light cone, that is,*

$$\widetilde{W}_n(\widetilde{f}_1,...,\widetilde{f}_n)=W_n(f_1,...,f_n)=0 \tag{B.2.1}$$

if there is a j such that $\widetilde{f}_j(p)=0$ for all $p\in\overline{V}_+$. This means $\widetilde{W}_n$ has its support inside the forward light cone $(\overline{V}_+)^n$.

(3) HERMITICITY. *$W_n(f_1,...,f_n)=W_n(f_n^*,...,f_1^*)^*$.*

(4) CAUSALITY. *If f_j and f_{j+1} have mutually spacelike separated support, then*

$$W_n(f_1,...,f_j,f_{j+1},...,f_n)=W_n(f_1,...,f_{j+1},f_j,...,f_n). \tag{B.2.2}$$

(5) POSITIVITY. *For all $f_{j,l}\in\mathscr{S}(\mathbb{M})$ one has*

$$\sum_{n\geq 0}\sum_{j+k=n}W_n(f_{j,j}^*,...,f_{j,1}^*,f_{k,1},...,f_{k,k})\geq 0, \tag{B.2.3}$$

where $W_0=|f_0|^2\geq 0$.

(6) CLUSTER DECOMPOSITION *Let $a\in\mathbb{M}$ be spacelike ($a^2<0$), then*

$$\lim_{\lambda\to\infty}W_n((f_1,...,f_j,\{\lambda a,1\}f_{j+1},...,\{\lambda a,1\}f_n)=W_j(f_1,...,f_j)W_{n-j}(f_{j+1},...,f_n). \tag{B.2.4}$$

Then there is a scalar field theory fulfilling the Wightman axioms 0 to IV. Any other theory is unitarily equivalent.

Proof. Axioms 0 & I: we start by considering the vector space $\mathfrak{D}$ given by terminating sequences

(B.2.5) $$\Psi_f = (f_0, f_1, f_2, ..., f_n, 0, 0, ...) \ , \qquad n \in \mathbb{N}$$

with elements $f_k \in \mathscr{S}(\mathbb{M}^k)$ and hence $\mathfrak{D} = \bigoplus_{n\geq 0} \mathscr{S}(\mathbb{M}^n)$, where $\mathscr{S}(\mathbb{M}^0) := \mathbb{C}$. We make use of the Wightman distributions to define an inner product on $\mathfrak{D}$ by

(B.2.6) $$\langle \Psi_f | \Psi_g \rangle = \langle (f_0, f_1, f_2, ...) | (g_0, g_1, g_2, ...) \rangle := \sum_{n\geq 0} \sum_{j+k=n} W_n(f_j^* \otimes g_k),$$

where for $n = 0$ we set $W_0(f_0, g_0) = f_0 g_0$, ie simply the product in $\mathbb{C}$. The sum in (B.2.6) is finite because the sequences terminate. $\langle \Psi_f | \Psi_g \rangle = \langle \Psi_g | \Psi_f \rangle^*$ is guaranteed by property (3). A representation of the Poincaré group is established by introducing the linear map

(B.2.7) $$U(a, \Lambda)(f_0, f_1, f_2, ...) := (f_0, \{a, \Lambda\} f_1, \{a, \Lambda\} f_2, ...),$$

where $(\{a, \Lambda\} f_j)(x_1, ..., x_n) := f_j(\Lambda^{-1}(x_1 - a), ..., \Lambda^{-1}(x_j - a))$, ie all arguments are equally transformed. Poincaré invariance of the inner product (B.2.6) is obvious from property (1). This implies in particular that the representation of the Poincaré group is unitary, ie

(B.2.8) $$\langle U(a, \Lambda)\Psi | U(a, \Lambda)\Psi' \rangle = \langle \Psi | \Psi' \rangle$$

holds for any states $\Psi = (f_0, f_1, f_2, ...)$, $\Psi' = (f_0', f_1', f_2', ...)$. Furthermore, we infer from (B.2.7) that the vacuum state is merely the vector

(B.2.9) $$\Psi_0 = (1, 0, 0, ...).$$

It is an easy exercise to show on states in $\mathfrak{D}$ that that this representation obeys the group law

(B.2.10) $$U(a', \Lambda')U(a, \Lambda) = U(a' + \Lambda' a, \Lambda' \Lambda)$$

and the property $||U(a, \Lambda)\Psi - \Psi|| \to 0$ as $(a, \Lambda) \to (0, 1)$. This latter properties implies strong continuity of the representation by the Cauchy-Schwartz inequality. The cluster decomposition property (6) implies that the vacuum is unique: let there be another Poincaré-invariant state $\Psi_0' \in \mathfrak{D}$. One may assume $\langle \Psi_0' | \Psi_0 \rangle = 0$. Then, for spacelike $a \in \mathbb{M}$, we have

(B.2.11)
$$\begin{aligned}\langle \Psi_0' | \Psi_0' \rangle &= \lim_{\lambda\to\infty} \langle \Psi_0' | U(\lambda a, 1)\Psi_0' \rangle = \lim_{\lambda\to\infty} \sum_{n\geq 0} \sum_{j+k=n} W_n(f_j^* \otimes U(\lambda a, 1) f_k) \\ &= \sum_{n\geq 0} \sum_{j+k=n} W_j(f_j^*) W_k(f_k) = \left(\sum_{n\geq 0} W_j(f_j^*) \right) \left(\sum_{m\geq 0} W_k(f_k) \right) = \langle \Psi_0' | \Psi_0 \rangle \langle \Psi_0 | \Psi_0' \rangle.\end{aligned}$$

We refer the interested reader to [**StreatWi00**] in which the completion of $\mathfrak{D}$ with respect to Cauchy series, nullifying of zero norm states and the spectral property of P^μ are discussed, the latter follows from property (2) .

Axiom II: a quantum field can now be defined through the assignment of $h \in \mathscr{S}(\mathbb{M})$ to the operation

(B.2.12) $$\varphi(h)(f_0, f_1, f_2, ...) = (0, f_0 h, h \otimes f_1, h \otimes f_2, ...).$$

All operators in the algebra $\mathcal{A}(\mathbb{M})$ are then of the form

(B.2.13) $$A = h_0 + \varphi(h_1) + \varphi(h_{2,1})\varphi(h_{2,2}) + ... + \varphi(h_{n,1})...\varphi(h_{n,n}) \in \mathcal{A}(\mathbb{M}).$$

Applying this to the vacuum generates the state $\Psi = (h_0, h_1, h_{2,1} \otimes h_{2,2}, ..., h_{n,1} \otimes ... \otimes h_{n,n}, 0, 0, ...)$ which is an element in $\mathfrak{D}$. It is obvious that this space is stable under the action of both the field and the Poincaré representation:

$$\varphi(\mathscr{S}(\mathbb{M}))\mathfrak{D} \subset \mathfrak{D}\ , \qquad U(\mathscr{P}_+^\uparrow)\mathfrak{D} \subset \mathfrak{D}. \tag{B.2.14}$$

However, the algebra $\mathcal{A}(\mathbb{M})$ generates only the dense subspace

$$\mathfrak{D}_0 := \mathcal{A}(\mathbb{M})\Psi_0 = \bigoplus_{n \geq 0} \mathscr{S}(\mathbb{M})^{\otimes n} \subset \mathfrak{D} \tag{B.2.15}$$

but not $\mathfrak{D}$. That the map $f \mapsto \langle \Psi | \varphi(f)\Psi' \rangle$ is a tempered distribution for all $\Psi, \Psi' \in \mathfrak{D}$ is obvious from

$$\langle \Psi | \varphi(f)\Psi' \rangle = \langle (h_0, h_1, h_2, ...) | \varphi(f)(g_0, g_1, g_2, ...) \rangle = \sum_{n \geq 0} \sum_{j+k=n} W_n(h_j \otimes f \otimes g_{k-1}) \tag{B.2.16}$$

because for all $n \in \mathbb{N}$ the assignment $f \mapsto W_n(h_j \otimes f \otimes g_{k-1})$ is a tempered distribution and the sum over all n is finite (we set $g_{-1} = 0$). Poincaré covariance (2.2.7) on $\mathfrak{D}$ and hence the validity of Axiom III is easy to see by

$$\begin{aligned} U(a,\Lambda)\varphi(h)(f_0, f_1, f_2, ...) &= U(a,\Lambda)(0, f_0 h, h \otimes f_1, h \otimes f_2, ...) \\ &= (0, f_0\{a,\Lambda\}h, \{a,\Lambda\}h \otimes \{a,\Lambda\}f_1, \{a,\Lambda\}h \otimes \{a,\Lambda\}f_2, ...) \\ &= \varphi(\{a,\Lambda\}h)(f_0, \{a,\Lambda\}f_1, \{a,\Lambda\}f_2, ...) \\ &= \varphi(\{a,\Lambda\}h)U(a,\Lambda)(f_0, f_1, f_2, ...), \end{aligned} \tag{B.2.17}$$

that is, $U(a,\Lambda)\varphi(h)U(a,\Lambda)^\dagger\mathfrak{D} = U(a,\Lambda)\varphi(h)U(a,\Lambda)^{-1}\mathfrak{D} = \varphi(\{a,\Lambda\}h)\mathfrak{D}$.

Local commutativity is guaranteed by property (4): let $A, B \in \mathcal{A}(\mathbb{M})$ generate the two states $\Psi = A\Psi_0$ and $\Phi = B\Psi_0$. If $f, g \in \mathscr{S}(\mathbb{M})$ have mutually spacelike support, then

$$\langle \Psi | \varphi(f)\varphi(g)\Phi \rangle = \langle \Psi_0 | B^*\varphi(f)\varphi(g)A\Psi_0 \rangle = \langle \Psi_0 | B^*\varphi(g)\varphi(f)A\Psi_0 \rangle = \langle \Psi | \varphi(g)\varphi(f)\Phi \rangle, \tag{B.2.18}$$

in which the second step makes use of property (4). Hence $[\varphi(f), \varphi(g)] = 0$ on $\mathfrak{D}_0$. Because $\mathfrak{D}_0 \subset \mathfrak{D}$ is dense this also holds on $\mathfrak{D}$.

To see that any other theory giving rise to the same Wightman distributions is unitarily equivalent to the one just constructed, let ϕ be the other field and Ω_0 its vacuum state. We define a linear map by setting

$$V\Psi_0 := \Omega_0\ , \qquad V\varphi(f_1)...\varphi(f_n)\Psi_0 := \phi(f_1)...\phi(f_n)\Omega_0. \tag{B.2.19}$$

Then for a general state $\Psi_f = (f_0, f_1, ..., f_n) \in \mathfrak{D}_0$ with $f_j = f_{j,1} \otimes ... \otimes f_{j,j} \in \mathscr{S}(\mathbb{M})^{\otimes j}$ we have

$$V\Psi_f = V[f_0\Psi_0 + \sum_{n \geq 1} \varphi(f_{n,1})...\varphi(f_{n,n})\Psi_0] = f_0\Omega_0 + \sum_{n \geq 1} \phi(f_{n,1})...\phi(f_{n,n})\Omega_0. \tag{B.2.20}$$

It is unitary because $\langle V\Psi_f | V\Psi_g \rangle = \langle \Psi_f | \Psi_g \rangle$ follows from coinciding vacuum expectation values. Next, we consider

$$\begin{aligned} V\varphi(h)\Psi_f &= V[f_0\varphi(h)\Psi_0 + \sum_{n \geq 1} \varphi(h)\varphi(f_{n,1})...\varphi(f_{n,n})\Psi_0] \\ &= f_0\phi(h)\Omega_0 + \sum_{n \geq 1} \phi(h)\phi(f_{n,1})...\phi(f_{n,n})\Omega_0 \\ &= \phi(h)[f_0\Omega_0 + \sum_{n \geq 1} \phi(f_{n,1})...\phi(f_{n,n})\Omega_0] = \phi(h)V\Psi_f \end{aligned} \tag{B.2.21}$$

which means $V\varphi(h) = \phi(h)V$ on $\mathfrak{D}_0$, ie $V\varphi(h)V^{-1} = \phi(h)$ on the dense subspace generated by the field ϕ. □

B.3. Jost-Schroer theorem

THEOREM B.3 (Jost-Schroer Theorem). *Let φ be a scalar field whose two-point Wightman distribution coincides with that of a free field with mass $m > 0$, ie*

$$\langle\Psi_0|\varphi(f)\varphi(h)\Psi_0\rangle = \int \frac{d^4p}{(2\pi)^4}\, \widetilde{f}^*(p)2\pi\theta(p_0)\delta(p^2-m^2)\widetilde{h}(p). \tag{B.3.1}$$

Then φ is itself a free field of the same mass.

PROOF. We follow [**StreatWi00**]. First note that because $\langle\Psi_0|\varphi([\Box+m^2]f)\varphi([\Box+m^2]h)\Psi_0\rangle = 0$, where $\Box + m^2$ is the Klein-Gordon operator, represented by the multiplication $\widetilde{f}(p) \mapsto (-p^2+m^2)\widetilde{f}(p)$ in momentum space for a Schwartz function f. This means in particular

$$||j(f)\Psi_0||^2 = \langle\Psi_0|\varphi([\Box+m^2]f^*)\varphi([\Box+m^2]f)\Psi_0\rangle = 0 \tag{B.3.2}$$

for the operator valued distribution $j(f) := \varphi([\Box+m^2]f)$ and thus $j(f)\Psi_0 = 0$. By the Reeh-Schlieder theorem (Theorem 2.7) we have $j(f) = 0$ for all $f \in \mathscr{S}(\mathbb{M})$. This means that φ obeys the free Klein-Gordon equation. We define the Fourier transform of the operator field by

$$\widetilde{\varphi}(\widetilde{f}) := \varphi(f), \tag{B.3.3}$$

where $\widetilde{f} \in \mathscr{S}(\mathbb{M})$ is the Fourier transform of $f \in \mathscr{S}(\mathbb{M})$ (and vice versa). The d'Alembertian takes the form of a multiplication operator

$$[\widetilde{\Box}+m^2]\widetilde{f}(p) = [-p^2+m^2]\widetilde{f}(p). \tag{B.3.4}$$

The Klein-Gordon equation for the Fourier transform of the field reads $\widetilde{\varphi}([\widetilde{\Box}+m^2]\widetilde{f}) = 0$ for all $\widetilde{f} \in \mathscr{S}(\mathbb{M})$. This means that $\widetilde{\varphi}$ vanishes unless the test function has some of its support inside the two hyperbolae

$$H_m^{\pm} = \{\ p \in M \mid p^2 = m^2, p_0 \gtrless 0\ \}. \tag{B.3.5}$$

The spectrum therefore is contained in the pair of mass hyperbolae, ie $\sigma(\widetilde{\varphi}) \subset H_m := H_m^+ \cup H_m^-$. We introduce two functions $\chi_\pm \in \mathscr{S}(\mathbb{M})$ with $\chi_\pm(p) = 1$ on the hyperbola $H_m^\pm$ and vanishing outside some neighbourhood of $H_m^\pm$. This enables us to split

$$\widetilde{\varphi}(\widetilde{f}) = \widetilde{\varphi}(\chi_-\widetilde{f}) + \widetilde{\varphi}(\chi_+\widetilde{f}) = \widetilde{\varphi}_+(\widetilde{f}) + \widetilde{\varphi}_-(\widetilde{f}), \tag{B.3.6}$$

where $\widetilde{\varphi}_\pm(\widetilde{f}) := \widetilde{\varphi}(\chi_\mp\widetilde{f})$, admittedly a very confusing convention. However, their Fourier transforms are

$$\varphi_\pm(f) := \widetilde{\varphi}_\pm(\widetilde{f}), \tag{B.3.7}$$

that is, the positive (negative) and negative (positive) frequency (energy) parts of the field. Because there are no negative energy states, one has

$$\varphi_+(f)\Psi_0 = 0. \tag{B.3.8}$$

Because the state $\varphi_+(f)\varphi_-(h)\Psi_0$ has a forward and a backward timelike momentum, its sum, the total momentum, is spacelike. One cannot create such a state. Either this state vanishes or it is a multiple of the vacuum. The only acceptable answer is that the latter is the case, since otherwise $\varphi_-(f)$ would annihilate the vacuum or be trivial, ie a multiple of the identity, neither of which makes sense. By (B.3.8) we see that

$$\langle\Psi_0|\varphi_+(f)\varphi_-(h)\Psi_0\rangle = \langle\Psi_0|\varphi(f)\varphi(h)\Psi_0\rangle = W(f,h) \tag{B.3.9}$$

and hence $[\varphi_+(f), \varphi_-(h)]\Psi_0 = \varphi_+(f)\varphi_-(h)\Psi_0 = W(f,h)\Psi_0$ because the state $\varphi_+(f)\varphi_-(h)\Psi_0$ is a multiple of the vacuum. The commutator then acts on the vacuum as

$$[\varphi(f), \varphi(h)]\Psi_0 = \{W(f,h) - W(h,f)\}\Psi_0 + [\varphi_-(f), \varphi_-(h)]\Psi_0. \tag{B.3.10}$$

The first thing we notice about the last term is $\langle \Psi_0 | [\varphi_-(f), \varphi_-(h)] \Psi_0 \rangle = 0$. The next aspect is that the distribution

$$F(f,h) := \langle \Psi | [\varphi_-(f), \varphi_-(h)] \Psi_0 \rangle \tag{B.3.11}$$

for any $\Psi \in \mathfrak{D}$ vanishes when f and h have mutually spacelike support. Consequently, the analytic continuation $F(z_1, z_2)$ into the forward tube $\mathcal{T}_2 = \mathbb{M}^2 + i(V_+)^2$ vanishes on the open subset of real spacelike $z_1 - z_2$. Since these vectors comprise an open set $E \subset \mathbb{M}^2$, this distribution vanishes by the edge of the wedge theorem (Theorem 2.5). Therefore, we have

$$([\varphi(f), \varphi(h)] - \{W(f,h) - W(h,f)\})\Psi_0 = 0, \tag{B.3.12}$$

ie the operator $T = [\varphi(f), \varphi(h)] - \{W(f,h) - W(h,f)\}$ annihilates the vacuum. Since by Theorem 2.7 no annihilator other than the null operator can be constructed from a locally generated polynomial algebra of fields, one has $T = 0$ and thus

$$[\varphi(f), \varphi(h)] = W(f,h) - W(h,f) \ , \qquad \varphi([\Box + m^2]f) = 0 \tag{B.3.13}$$

for all $f, h \mathscr{S}(\mathbb{M})$. This entails $W_n = 0$ for n odd and

$$W_n(f_1, ..., f_n) = \sum_{(i,j)} W(f_{i_1}, f_{j_1}) ... W(f_{i_m}, f_{j_m}) \tag{B.3.14}$$

for even $n = 2m$, where the sum is over all permutations of $1, 2, ..., n$ written as $i_1, j_1, ..., i_m, j_m$ such that $i_1 < i_2 < ... < i_m < j_1 < ... < j_m$ (Wick contractions). □

B.4. DSE for the anomalous dimension of the photon

We now take the Dyson-Schwinger equation for the Green's function of the photon

$$G^{\sim}(\alpha, L) = 1 - \sum_{j \geq 1} \alpha^j \, G^{\sim}(\alpha, -\partial_\rho)^{1-j} [e^{-\rho L} - 1] H_j(\rho) \big|_{\rho = 0} \tag{B.4.1}$$

and derive in the following the DSE for its anomalous dimension, ie equation (5.5.21). In a first step we differentiate this equation with respect to L and then set $L = 0$ to get

$$\gamma_1^{\sim}(\alpha) = \sum_{j \geq 1} \alpha^j \, [1 - \gamma^{\sim}(\alpha) \cdot (-\partial_\rho)]^{1-j} (-\rho) H_j(\rho) \big|_{\rho = 0} \ , \tag{B.4.2}$$

where $\gamma^{\sim}(\alpha) \cdot (-\partial_\rho) = \sum_{r \geq 1} \gamma_r^{\sim}(\alpha)(-\partial_\rho)^r$ is a shorthand. We write it now as U and aim to compute

$$\gamma^{\sim}(\alpha) = \sum_{j \geq 1} \alpha^j \lim_{\rho \to 0} (1 - U)^{1-j} (-\rho) H_j(\rho) = \alpha \lim_{\rho \to 0} (-\rho) H_1(\rho) + \sum_{j \geq 2} \alpha^j \lim_{\rho \to 0} (1 - U)^{1-j} (-\rho) H_j(\rho). \tag{B.4.3}$$

Inserting the Laurent series $H_j(\rho) = \sum_{t \geq 0} h_t^{(j)} \rho^{t-1}$ yields

$$\begin{aligned} \gamma^{\sim}(\alpha) &= -\alpha h_0^{(1)} - \sum_{t \geq 0} \sum_{j \geq 2} \alpha^j h_t^{(j)} \lim_{\rho \to 0} (1 - U)^{1-j} \rho^t \\ &= J(\alpha) - \sum_{t \geq 1} \sum_{j \geq 2} \alpha^j h_t^{(j)} \lim_{\rho \to 0} [1 + \sum_{m \geq 1} \binom{1-j}{m} (-U)^m] \rho^t, \end{aligned} \tag{B.4.4}$$

where $J(\alpha) = -\sum_{j \geq 1} h_0^{(j)} \alpha^j$. We first find

$$\lim_{\rho \to 0} (-U)^m \rho^t = \lim_{\rho \to 0} [-\sum_{r \geq 1} (-1)^r \gamma_r^{\sim}(\alpha) \partial_\rho^r]^m \rho^t = \sum_{r \geq m} (-1)^{r+m} [\sum_{r_1 + ... + r_m = r} \gamma_{r_1}^{\sim}(\alpha) ... \gamma_{r_m}^{\sim}(\alpha) \ ] \ r! \delta_{t,r}. \tag{B.4.5}$$

Then plugging this into (B.4.6) gives the result

$$\gamma^{\sim}(\alpha) = J(\alpha) + \sum_{m\geq 1}\sum_{r\geq m} J_{m,r}(\alpha) r!\ [\sum_{r_1+\ldots+r_m=r} \gamma^{\sim}_{r_1}(\alpha)\ldots\gamma^{\sim}_{r_m}(\alpha)\ ] \tag{B.4.6}$$

with $J_{m,r}(\alpha) = (-1)^{r+m+1}\sum_{j\geq 2}\binom{1-j}{m} h_r^{(j)}\alpha^j$.

List of frequently used abbreviations and symbols

$\langle S\rangle_{\mathbb{F}} = \mathbb{F}[S]$	freely generated commutative polynomial $\mathbb{F}$-algebra based on a set S
$\mathscr{A}[[\mathfrak{M}]]$	algebra of transseries with trans coefficients in the algebra $\mathscr{A}$
$\mathcal{A}(X)$	(noncommutative) algebra generated by field operators $\varphi(f)$, $f \in \mathscr{D}(X)$
$\mathscr{B}f$	Borel transform of a formal power series f
$\mathrm{ch}(\mathcal{H},\mathbb{R})$	Lie algebra of infinitesimal Hopf algebra characters from $\mathcal{H}$ to $\mathbb{R}$
$\mathrm{Ch}(\mathcal{H},\mathcal{A})$	group of Hopf algebra characters from $\mathcal{H}$ to $\mathcal{A}$
χ, χ_R, S_R^{χ}	Hopf algebra character, renormalised cousin and counterterm character
DSE, DSEs	Dyson-Schwinger equation, Dyson-Schwinger equations
$\mathscr{D}(X)$	set of Schwartz functions with compact support in $X \subseteq \mathbb{R}^n$.
$\mathfrak{D}, \mathfrak{D}_0$	dense subspaces of Hilbert space $\mathfrak{H}$, $\mathfrak{D}_0$ field algebra
$\varphi, \varphi_r, \varphi_0$	(bare) scalar field and renormalised scalar field, free field
$\gamma_j(z)$	j-th RG (or log-coefficient) function with coupling z
$\widetilde{\gamma}_j(\mathfrak{m})$	j-th RG transseries with real coefficients
$\overline{\gamma}_j(\mathfrak{m})$	j-th RG transseries with coefficients in $\mathscr{A}$
$\mathcal{H}$	Hopf algebra of Feynman graphs (for whatever renormalisable theory)
$\mathfrak{H}, \mathfrak{H}_0$	Hilbert space , Hilbert space of the free theory
H, H_0, H_{int}	interacting Hamiltonian, free Hamiltonian, interaction part of H
$\mathscr{H}_I(x)$	interaction picture Hamiltonian, ie interaction part in Dirac picture
$\mathscr{L}_+^{\uparrow}$	proper orthochronous Lorentz group
$\mathscr{L}, \mathscr{L}(\varphi)$	Lagrangian (Lagrangian density), Lagrangian of scalar field φ
$\mathscr{L}_r, \mathscr{L}_{ct}$	renormalised Lagrangian, counterterm Lagrangian
$\mathscr{L}_{int}$	interaction part of Lagrangian
$\mathscr{L}[\mathscr{B}P](\alpha)$	Borel sum of $P(\alpha)$ (Borel-Laplace transform of Borel transform)
$\mathcal{L}(\mathcal{H})$	vector space of linear maps on the Hopf algebra $\mathcal{H}$
$\mathbb{M}$	Minkowski space $\mathbb{R}^4$
$\mathfrak{M}, \mathfrak{m}$	set of atomic transmonomials, transmonomial triple $\mathfrak{m} = (\mathfrak{m}_1, \mathfrak{m}_2, \mathfrak{m}_3)$
$\mathscr{P}_+^{\uparrow}$	connected Poincaré group
RG	renormalisation group
R_θ	RG (renormalisation group) recursion operator
$\mathscr{R}$	renormalisation map (for differential forms)
$\mathbb{R}[[\mathfrak{M}]]$	algebra of transseries with real trans coefficients
$\mathscr{S}(\mathbb{M})$	the algebra of Schwartz functions (fast decreasing and smooth)
T	time-ordering operator
$\mathcal{T}_n, \mathcal{T}_n' \subset \mathbb{C}^{4n}$	forward tube, extended forward tube
$\mathbb{T}_1(\mathfrak{M}), \mathbb{T}_2(\mathfrak{M})$	instanton-, loop-homogeneous transseries
$U(a,\Lambda)$	strongly continuous unitary representation of $\mathscr{P}_+^{\uparrow}$
V	unitary intertwiner between two Hilbert spaces

Bibliography

[AlHoMa08] S.A. Albeverio, R.J. Høegh-Krohn, S. Mazzucchi : *Mathematical Theory of Feynman Path Integrals. An introduction.* Springer (2008)

[AnSchi15] I. Aniceto, R. Schiappa : *Nonperturbative Ambiguities and the Reality of Resurgent Transseries.* Comm. Math. Phys. 335, 183-245 (2015), arXiv: hep-th/1308.1115

[AMY12] K.V. Antipin, M.N. Mnatsakanova, Yu.S. Vernov: *Haag's theorem in noncommutative quantum field theory.* arXiv: math-ph/1202.0995v1

[Bar63] G. Barton: *Introduction to Advanced Field Theory.* Wiley Interscience (1963)

[Bau87] K. Baumann: *On relativistic irreducible quantum fields fulfilling CCR.* J. Math. Phys. 28, 697-704 (1987)

[Bau88] K. Baumann: *On canonical irreducible quantum field theories describing bosons and fermions.* J. Math. Phys. 29, 1225-1230 (1988)

[BeCla14] M.P. Bellon, P. J. Clavier : *Higher order corrections to the asymptotic perturbative solution of a Schwinger-Dyson equation.* Lett. Math. Phys. 104, 1-22 (2014), arXiv: hep-th/1311.1160v2

[Ben99] M. Beneke: *Renormalons*, Phys. Rep. 317 (1999)

[BerKrei06] C. Bergbauer, D. Kreimer : *Hopf algebras in renormalisation theory: locality and Dyson-Schwinger equations from Hochschild-Cohomology.* IRMA Lect. Math. Theor. Phys. 10, 133-164 (2006), arXiv: hep-th/0506190

[BiKreiW07] I.Bierenbaum, D.Kreimer, S.Weinzierl: *The next-to-ladder approximation for linear Dyson-Schwinger equations*, Phys. Lett. B 646 (2007), 129-133, arXiv: hep-th/0612180v2

[BKUY09] G. van Baalen, D.Kreimer, D.Uminsky, K.Yeats: *The QED β-function from global solutions to Dyson-Schwinger equations*, Ann. Phys. 324 (2009), 205-219, arXiv: hep-th/0805.0826

[BKUY10] G. van Baalen, D.Kreimer, D.Uminsky, K.Yeats: *The QCD β-function from global solutions to Dyson-Schwinger equations*, Ann. Phys. 325 (2010), 205-219, arXiv: hep-th/0805.0826

[Blee81] D.Bleecker: *Gauge Theory and Variational Principles.* Mineola, New York (1981)

[Bleu50] K. Bleuler : *Eine neue Methode zur Behandlung der longitudinalen und skalaren Photonen.* Helv. Phys. Act. 23, 567-586 (1950)

[BjoDre65] J.D. Bjorken, S.D. Drell : *Relativistic Quantum Fields.* McGraw-Hill (1965)

[Bogo75] N.N. Bogoliubov, A.A. Logunov, I.T. Todorov: *Introduction to Axiomatic Quantum Field Theory.* W.A. Benjamin, Inc. (1975)

[Bor62] H.-J. Borchers: *On Structure of the algebra of field operators.* Nuovo Cimento 24, 214- 236 (1962)

[BoHaSch63] H.J. Borchers, R. Haag, B. Schroer : *The Vacuum State in Quantum Field Theory.* Nuovo Cim. 29, 148 (1963)

[BoRo33] N. Bohr, L. Rosenfeld : *Zur Frage der Messbarkeit der elektromagnetischen Feldgrössen.* Danske Videnskabernes Selskab Matematisk-Fysiske Meddelelser 12, 1-65 (1933)

[BoRo50] N. Bohr, L. Rosenfeld : *Field and Charge Measurementa in Quantum Electrodynamics.* Phys. Rev. 78, 794-798 (1950)

[BroK01] D.J. Broadhurst, D. Kreimer: *Exact solution of Dyson-Schwinger equations for iterated one-loop integrals and propagator coupling duality*, Nucl. Phys. B 600 (2001), 403 - 422, arXiv: hep-th/0012146

[BrowKr11] F.Brown, D.Kreimer: *Angles, Scales and Parametric Renormalization*, Lett. Math. Phys. 103, 933-1007 (2013), arXiv: hep-th/1112.1180

[BruFK96] R. Brunetti, K. Fredenhagen, M. Köhler: *The microlocal spectrum condition and the Wick polynomials of free fields.* Comm. Math. Phys. 180, 633 - 652 (1996)

[Bu75] D. Buchholz: *Collision theory for massless Fermions.* Comm. Math. Phys. 42, 269-279 (1975)

[Bu77] D. Buchholz: *Collision theory for massless Bosons.* Comm. Math. Phys. 52, 147-173 (1977)

[Bu00] D. Buchholz: *Algebraic Quantum Field Theory: A Status Report.* arXiv: math-ph/0011044v1 (2000)

[BuHa00] D. Buchholz, R. Haag: *The Quest for Understanding in Relativistic Quantum Physics.* J. Math. Phys. 41, 3674-3697 (2000), arXiv: hep-th/9910243

[Chai68] J.M. Chaiken : *Number Operators for Representations of the Canonical Commutation Relations.* Comm. Math. Phys. 8, 164-184 (1968)

[Co53] J.M. Cook: *The Mathematics of Second Quantization.* Trans. Amer. Math. Soc. 74, 222-245 (1953)

[CoKrei98] A. Connes, D. Kreimer : *Hopf algebras, renormalisation and noncommutative geometry.* Comm. Math. Phys. 199, 203-242 (1998), arXiv: hep-th/9808042

[CoKrei00] A. Connes, D. Kreimer: *Renormalisation in quantum field theory and the Riemann-Hilbert problem I: The Hopf algebra structure of graphs and the main theorem.* Comm. Math. Phys. 210, 249-273 (2000)

[CoSaESVo15] R. Couso-Santamaria, J. D. Edelstein, R. Schiappa, M. Vonk: *Resurgent transseries and the Holomorphic Anomaly.* arXiv: hep-th/1308.1695v3

[DeDoRu66] G.-F. Dell'Antonio, S. Doplicher, D. Ruelle : *A Theorem on Canonical Commutation Relations.* Comm. Math. Phys. 2, 223 -230 (1966)

[DeDo67] G.-F. Dell'Antonio, S. Doplicher : *Total Number of Particles and Fock representation.* J. Math. Phys. 8, 663-666 (1967)

[DeKaTh96] R. Delbourgo, A.C. Kalloniatis, G. Thompson : *Dimensional renormalisation: ladders and rainbows.* Phys. Rev. D 54, 5373-5376 (1996)

[DelEM97] R. Delbourgo, D. Elliott, D.S. McAlly: *Dimensional Renormalization in ϕ^3 theory: Ladders and rainbows*, Phys. Rev. D (1997), 5230

[DeKaTh97] R. Delbourgo, A.C. Kalloniatis, G. Thompson : *Dimensional renormalisation: ladders and rainbows.* Phys. Rev. D 55, 5230-5233 (1997)

[Di73] J. Dimock: *Asymptotic Perturbation Expansion in the $P(\phi)_2$ Quantum Field Theory.* Comm. Math. Phys. 35, 347-356 (1974)

[Di11] J. Dimock : *Quantum Mechanics and Quantum Field Theory. A Mathematical Primer.* Cambridge (2011)

[Dix58] J.Dixmier: *Sur la relation $i(PQ - QP) = 1$.* Comp. Math. 13, 263-270 (1958)

[Dori14] D. Dorigoni : *An Introduction to Resurgence, Trans-Series and Alien Calculus.* arXiv: hep-th/1411.3585v1

[DresKa62] M. Dresden, P.B. Kahn: *Field theories with persistent one particle states. I. General formalism.* Rev. Mod. Phys. 34, 401 (1962)

[Dunc21] A. Duncan: *The Conceptual Framework of Quantum Field Theory.* Oxford University Press (2012)

[DunUen12] G.V. Dunne, M. Ünsal : *Resurgence and trans-series in Quantum Field Theory: the $\mathbb{CP}^{N-1}$ model.* JHEP 11, 170 (2012), arXiv: hep-th/1210.2423

[DunUen13] G.V. Dunne, M. Ünsal : *Continuity and Resurgence: Towards a continuum definition of the $\mathbb{CP}^{N-1}$ model.* Phys. Rev. D 87, 025015 (2013)

[DunUen14] G. V. Dunne, M. Ünsal: *Generating nonperturbative physics from perturbation theory.* Phys. Rev. D 89, 041701 (2014) arXiv: hep-th/1306.4405v2 (2014)

[Dys49a] F.J. Dyson: *The Radiation Theories of Tomonaga, Schwinger, and Feynman.* Phys. Rev. 75, 486 - 502 (1949)

[Dys49b] F.J. Dyson: *The S Matrix in Quantum Electrodynamics.* Phys. Rev. 75, 1736 -1755 (1949)

[Dys51] F. J. Dyson : *Divergence of Perturbation Theory in Quantum Electrodynamics*, Phys. Rev. 85 (1951), 631

[EaFra06] J. Earman, D. Fraser: *Haag's theorem and its implication for the foundations of quantum field theory.* Erkenntnis, 64, 305-344 (2006)

[Eca81] J. Ecalle's : *Les fonctions resurgentes.* Publ. Math. d'Orsay, Vol.I-III (1981)

[Ed09] G.A. Edgar: *Transseries for beginners.* Real Anal. Exchange 35, 253-310 (2009), arXiv: math.RA/0801.4877v5

[EGraPa07] K.Ebrahimi-Fard, J. Gracia-Bondia, F. Patras: *A Lie theoretic approach to renormalisation.* Comm. Math. Phys. 276, 519-549 (2007), arXiv: hep-th/0609035

[Em09] G.G.Emch: *Algebraic Methods in Statistical Mechanics and Quantum Field Theory.* Dover Publications, Inc. (2009)

[FaSi97] S.V. Faleev, P.G. Silvestrov : *Higher order corrections to the Renormalon.* arXiv: hep-th/9610344v2

[FePStro74] F. Strocchi : *Some Remarks on Local Operators in Quantum Electrodynamics.* Comm. Math. Phys. 35, 25-38 (1974)

[FeJo60] P.G. Federbush, K.A. Johnson: *Uniqueness Properties of the Twofold Vacuum Expectation Value.* Phys. Rev. 120, 1926 (1960)

[Feyn49] R.P. Feynman: *Space-Time Approach to Quantum Electrodynamics.* Phys. Rev. 49, 769 -789 (1949)

[Feyn06] R.P. Feynman : *QED. The strange theory of light and matter.* Princeton University Press (2006)

[Fo32] V.Fock: *Konfigurationsraum und zweite Quantelung.* Zeitschr. f. Phys. 75, 622-647 (1932)

[Foi10] L. Foissy : *Classification of systems of Dyson-Schwinger equations in the Hopf algebra of decorated rooted trees.* Adv. Math. 224, 2094-2150 (2010)

[Fra06] D.L. Fraser: *Haag's Theorem and the interpretation of quantum field theories with interactions.* Thesis, University of Pittsburgh (2006)

[Fred10] K. Fredenhagen: *Quantum field theory.* Lecture notes, Universität Hamburg (2010), available online at `unith.desy.de/research/aqft/lecture_notes`

[Fried53] K.O.Friedrichs: *Mathematical Aspects of the Quantum Theory of Fields.* Interscience Publishers, Inc., New York (1953)

[Fro82] J. Fröhlich: *On the triviality of $\lambda\varphi_d^4$ theories and the approach to the critical point in $d \geq 4$ dimensions.* Nucl. Phys. B 200, 281-296 (1982)

[GaKaG98] E.Gardi, M.Karliner & G.Grunberg: *Can the QCD running coupling have a causal analyticity structure?*, JHEP 07 (1998), 007

[GaWi54] L.Gårding, A. Wightman: *Representations of the Commutation Relations.* Proc. Nat. Acad. Sci. 40, 622-626 (1954)

[GeMLo51] M. Gell-Mann, F. Low: *Bound States in Quantum Field Theory.* Phys. Rev. 84, 350-354 (1951)

[GeMLo54] M. Gell-Mann, F. Low: *Quantum Electrodynamics at Small Distances States.* Phys. Rev. 95, 1300 (1954)

[GliJaf68] J.Glimm, A.Jaffe: *A $\lambda\phi^4$ quantum theory without cutoffs I.* Phys. Rev. 176, 1945-1951 (1968)

[GliJaf70] J.Glimm, A.Jaffe: *A $\lambda\phi^4$ quantum theory without cutoffs II.* Ann. Math. 91, 362-401 (1970)

[GliJaf81] J.Glimm, A.Jaffe: *Quantum Physics. A Functional Integral Point of View.* Springer (1981) Ann. Math. 91, 362-401 (1970)

[GliJaSp74] J. Glimm, A. Jaffe, T. Spencer: *The Wightman axioms and particle structure in the $P(\phi)_2$ quantum field model.* Ann. Math. Phys. 100, 585-632 (1974)

[GoKaLaSu91] S.G.Gorishny, A.L.Kataev, S.A.Larin, L.R.Surguladze: *The analytic four-loop corrections to the QED β-function in the MS scheme and to the QED ψ-function. Total reevaluation.* Phys. Lett. B 256 (1991), 81

[GraMaZa15] A. Grassi, M. Marino, S. Zakany: *Resumming the perturbation series.* arXiv: hep-th/1405.4214v2

[Gre59] O.W. Greenberg: *Haag's Theorem and Clothed Operators.* Phys. Rev. 115, 706-710 (1959)

[Gre61] O.W. Greenberg : *Generalized Free Field and Models of Local Field Theory.* Ann. Phys. 16, 158-176 (1961)

[Gu50] S. N. Gupta :*Theory of Longitudinal Photons in Quantum Electrodynamics.* Proc. Phys. Soc. 63, 681-691 (1950)

[Gue66] M. Guenin : *On the Interaction Picture.* Comm. math. phys. 3, 120 - 132 (1966)

[GueRoSi72] F. Guerra, L. Rosen, B. Simon: *Nelson's Symmetry and the infinite Volume Behaviour of the vacuum in $P(\phi)_2$.* Comm. Math. Phys. 27, 10-22 (1972)

[Ha55] R.Haag: *On Quantum Field Theories.* Danske Videnskabernes Selskab Matematisk-Fysiske Meddelelser 29, 1-37 (1955)

[Ha58] R. Haag: *Quantum Field Theories with Composite Particles and Asymptotic Conditions.* Phys. Rev. 112, 669-673 (1958)

[Ha96] R. Haag: *Local Quantum Physics.* Springer (1996)

[HaKa64] R.Haag, D.Kasterl: *An Algebraic Approach to Quantum Field Theory.* J. Math. Phys. 5, 848-861 (1964)

[HeiPau29] W.Heisenberg, W.Pauli: *Zur Quantendynamik der Wellenfelder.* Zeitschr. f. Phys. 56, 1-61 (1929)

[He63] K. Hepp: *Lorentz-kovariante analytische Funktionen.* Helv. Phys. Acta 36, 355 (1963)

[He66] K. Hepp: *Proof of the Bogoliubov-Parasiuk Theorem on Renormalisation.* Comm. Math. Phys. 2, 301-326 (1966)

[Hoe06] J.van der Hoeven: *Transseries and real Differential Algebra.* Springer (2006)

[ItZu80] C. Itzykson, J.-B. Zuber: *Quantum Field Theory.* Dover Publications, Inc. (1980)

[Jaff69] A. Jaffe: *Whither axiomatic quantum field theory?* Rev. Mod. Pyhs. 41, 576 - 580 (1969)

[JenZin04] U- D. Jentschura, J. Zinn-Justin: *Instantons in quantum mechanics and resurgent expansions.* Phys. Lett. B 596, 138-144 (2004)

[Jo61] R. Jost: *Properties of Wightman Functions.* In: Lectures on Field Theory and the Many-Body Problem, edited by E.R. Caianiello, Academic Press (1961)

[Jo65] R. Jost: *The General Theory of Quantized Fields.* Amercian Mathematical Society, (1965)

[Jo08] J. Jost: *Dynamical Systems.* Springer (2008)

[Ka12] C. Kassel : *Quantum groups.* Springer (2012)

[Ki12] H. Kißler : *On linear systems of Dyson-Schwinger equations.* Master's Thesis, Humboldt University Berlin (2012)

[KlaKrei13] L.Klaczynski, D.Kreimer: *Avoidance of a Landau pole by flat contributions in QED.* Ann. Phys. 344, 213-231 (2014), arXiv: hep-th/1309.5061

[KlaKrei13a] L.Klaczynski, D.Kreimer : *Renormalization & Renormalization Group.* Lecture notes; lectures by D. Kreimer, winter term 2012/2013 (2013)

[KlaKrei13b] L.Klaczynski, D.Kreimer : *Dyson-Schwinger equations.* Lecture notes; lectures by D. Kreimer, summer term 2013 (2013)

[Krei99] D. Kreimer : *On overlapping divergences.* Comm. Math. Phys. 204, 669 (1999), arXiv: hep-th/9810022

[Krei02] D. Kreimer: *Combinatorics of (perturbative) quantum field theory.* Phys. Rep. 363, 387-424 (2002)

[Krei06] D. Kreimer: *Etude for linear Dyson-Schwinger Equations*, IHES/P/06/23 (see internet)

[Krei06a] D. Kreimer : *Anatomy of a gauge theory.* Ann. Phys. 321, 2757-2781 (2006), arXiv: hep-th/0509135

[KreiPa12] D.Kreimer, E. Panzer: *Renormalisation and Mellin transforms.* Computer Algebra in Quantum Field Theory, Springer (2013) arXiv: hep-th/1207.6321

[KrY06] D.Kreimer, K.Yeats: *An étude in non-linear Dyson-Schwinger equations*, Nucl. Phys. B Proc. Supp., 160, 116-121 (2006), arXiv: hep-th/0605096

[LeBlo67] J.-M. Lévy-Leblond: *Galilean Quantum Field Theories and a Ghostless Lee Model.* Comm. Math. Phys. 4, 157-176 (1967)

[Lo61] J. Lopuszanski : *A Criterion for the Free Field Character of Fields.* J. Math. Phys. 2, 743 - 747 (1961)

[LSZ55] H.Lehmann, K.Symanzik, W.Zimmermann: *Zur Formulierung quantisierter Feldtheorien.* Il Nuovo Cimento 1, 205-225 (1955)

[LSZ57] H.Lehmann, K.Symanzik, W.Zimmermann: *On the Formulation of Quantized Field Theories II.* Il Nuovo Cimento 6, 319-333 (1957)

[Lu05] T. Lupher: *Who proved Haag's Theorem?* International Journal of Theoretical Physics 44, 1995 - 2005 (2005)

[Man04] D. Manchon : *Hopf algebras, from basics to applications to renormalisation.* arXiv: math/0408405

[Mar14] M. Marino : *Lectures on nonperturbative effects in large N gauge theories, matrix models and strings.* arXiv: hep-th/1206.6272v2

[Nest03] A.V. Nesterenko : *Analytic invariant charge in QCD*, J.Mod.Phys. A 18, No.30 (2003), 5475, archiv: hep-th/0308288v2

[OS76] K. Osterwalder, R. Sénéor: *The scattering matrix is nontrivial for weakly coupled $P(\varphi)_2$ models.* Helv. Phys. Act. 49, 525-535 (1976)

[OSchra73] K.Osterwalder, R.Schrader: *Axioms for Euclidean Green's Functions.* Comm. Math. Phys. 31, 83-112 (1973)

[PenRo81] M.R. Pennington, G.G. Ross: *Perturbative QCD for timelike processes: What is the best expansion parameter?*, Phys. Lett. B 102 (1981), 167

[PeSch95] M.E. Peskin, D.V. Schroeder: *An Introduction to Quantum Field Theory.* Westview Press (1995)

[Po69] K. Pohlmeyer : *The Jost-Schroer Theorem for Zero-Mass Fields.* Comm. Math. Phys. 12, 204-211 (1969)

[Pow67] R.T. Powers : *Absence of Interactions as a Consequence of Good Ultraviolet Behaviour in the Case of a Local Fermi Field.* Comm. Math. Phys. 4, 145-156 (1967)

[Ree88] H.Reeh: *A remark concerning commutation relations.* J.Math.Phys. 29 (7), 1535 (1988)

[ReSi75] M. Reed, B. Simon : *Methods of Modern Mathematical Physics.* Vol. 2: Fourier Analysis, Self-Adjointness, Academic Press (1975)

[Ri14] D. Rickles: *A Brief History of String Theory.* Springer (2014)

[Ro69] P. Roman: *Introduction to Quantum Field Theory.* John Wiley and Sons, Inc. (1969)

[Rue11] L.Ruetsche: *Interpretating Quantum Theories.* Oxford University Press (2011), p.18

[Rue62] D. Ruelle: *On the Asymptotic Condition in Quantum Field Theory.* Helv. Phys. Acta, 35, 147 (1962)

[Sa07] D. Sauzin: *Resurgent functions and splitting problems.* arXiv: math.DS/0706.0137v1

[Sa14] D. Sauzin: *Introduction to 1-summability and resurgence.* arXiv: math.DS/1405.0356

[Scha14] G. Scharf: *Finite Quantum ELectrodynamics. The Causal Approach.* Dover Publications, Ic. (2014)

[Scho08] M. Schottenloher: *A Mathematical Introduction to Conformal Field Theory.* Springer, 2008

[Schra74] R. Schrader: *On the Euclidean Version of Haag's Theorem in $P(\varphi)_2$ field theories.* Comm. Math. Phys. 36, 133-136 (1974)

[Schra76] R. Schrader: *A possible Constructive Approach to $(\phi^4)_4$.* Comm. Math. Phys. 49, 131-153 (1976)

[Schw94] S.Schweber: *QED and the men who made it.* Princeton University Press (1994)

[Schwi51] J. Schwinger : *On the Green's functions of quantized fields. I & II.* Proc. Nat. Acad. Sc. 37, 452-459 (1951)

[Seg67] I. Segal: *Notes toward the construction of non-linear relativistic quantum fields I. The Hamiltonian in two space-time dimensions as the generator of a C*-automorphism group.* Proc. Nat. Am. Soc. 57, 1178 - 1183 (1967)

[SinEm69] K.Sinha, G.G. Emch: *Adaptation of Powers' No-Interaction Theorem to Bose Fields.* Bull. Am. Phys. Soc. 14, 86 (1969)

[Sok79] A.D. Sokal: *An improvement of Watson's theorem on Borel summability.* J. Math. Phys. 21, 261-263 (1979)

[Sta62] H.P. Stapp: *Axiomatic S-Matrix Theory.* Rev. Mod. Phys. 34, 390-394 (1962)

[Stei00] O. Steinmann: *Perturbative Quantum Electrodynamics and Axiomatic Field Theory.* Springer (2000)

[Ster93] G. Sterman : *An introduction to Quantum Field Theory.* Cambridge University Press (1993)

[Sti02] M. Stingl: *Field-Theory Amplitudes as Resurgent Functions.* arXiv: hep-ph/0207349

[Strau01] N.Straumann: *Schrödingers Entdeckung der Wellenmechanik.* arXiv: quant-ph/0110097v1 (2001)

[Strau13] N.Straumann: *Quantenmechanik.* Springer Verlag (2013)

[Streat75] R.F.Streater: *Outline of axiomatic quantum field theory.* Rep. Prog. Phys. 38, 771-846 (1975)

[StreatWi00] R.F.Streater, A.S.Wightman: *PCT, Spin and Statistics, and all that.* Princeton University Press (2000)

[Strei68] L. Streit : *A Generalization of Haag's Theorem.* Il Nuovo Cimento 72, 673-680 (1969)

[Stro67] F. Strocchi : *Gauge Problem in Quantum Field Theory.* Phys. Rev. 162, 1429-1438 (1967)

[Stro70] F. Strocchi : *Gauge Problem in Quantum Field Theory III. Quantization of Maxwell Equations and Weak Local Commutativity.* Phys. Rev. 162, 2334-2340 (1970)

[Stro93] F.Strocchi: *Selected topics on the general properties of quantum field theory.* World Scientific (1993)

[Stro13] F.Strocchi: *An Introduction to Non-Perturbative Foundations of Quantum Field Theory.* Oxford (2013)

[StroWi74] F.Strocchi, A.S.Wightman: *Proof of the charge superselection rule in local relativistic quantum field theory.* J. Math. Phys. 15, 2198-2224 (1974)

[Stue50] E.C.G. Stueckelberg: *Relativistic Quantum Theory for Finite Time Intervals.* Phys. Rev. 81, 130-133 (1950)

[Su01] S.J.Summers: *On the Stone-von Neumann Theorem and its Ramifications*, in: *John von Neumann and the Foundations of Quantum Physics*, edited by M.Rédei, M.Stöltzner, Kluwer Academic Publishers (2001)

[Su12] S.J.Summers: *A Perspective on Constructive Quantum Field Theory.* arXiv: math-ph/1203.3991v1 (2012)

[Sui06] W.D. van Suijlekom: *The Hopf Algebra of Feynman Graphs in Quantum Electrodynamics.* Letters in Mathematical Physics 77, 265-281 (2006)

[Sui07] W.D. Suijlekom: *Multiplicative renormalisation and Hopf algebras.* hep-th/0707.0555

[Sw69] M.E. Sweedler: *Hopf algebras.* W. A. Benjamin, Inc., New York (1969)

[TaTaYa80] Y. Taguchi, A. Tanaka, K. Yamamoto: *Haag's Theorem and dimensional regularisation.* Phys. Letters 90 B, 422-423 (1980)

[Te12] G. Teschl: *Ordinary Differential Equations and Dynamical Systems.* AMS Vol. 140, Graduate Studies in Mathematics (2012)

[Ti99] Ticciati, R. : *Quantum Field Theory for Mathematicians.* Cambridge University Press (1999)

[vHo52] L. van Hove: *Les difficultés de divergences pour un modelle particulier de champ quantifié.* Physica 18, 145-159 (1952)

[Wa50] J.C. Ward : *An identity in Quantum Electrodynamics.* Phys. Rev. 78, 182 (1950)

[Wei11] M. Weier : *An Algebraic Version of Haag's Theorem.* Comm. Math. Phys. 305, 469-485 (2011)

[vNeu31] J.von Neumann: *Zur Eindeutigkeit der Schrödingerschen Operatoren.* Ann. Math. 104, 570-578 (1931)

[Wein95] S. Weinberg: *The Quantum Theory of Fields.* Vols.I-III, Cambridge University Press (1995)

[Wic50] G.C. Wick: *The Evaluation of the Collision Matrix.* Phys. Rev. 80, 268 (1950)

[Wi56] A. Wightman : *Quantum Field Theory in Terms of Vacuum Expectation Values.* Phys. Rev. 56, 860-866 (1956)

[Wi64] A. Wightman: *La théorie quantique locale et la théorie quantique des champs.* Ann. Inst. H. Poincaré I, 403-420 (1964)

[Wi67] A. Wightman : *Introduction to some aspects of the relativistic dynamics of quantised fields.* In: M. Lévy: *Cargése Lectures in Theoretical Physics: High Energy Interactions and Field Theory.* Gordon and Breach, pp.171-291 (1967)

[WiHa57] D. Hall, A. Wightman: *A theorem on invariant analytic functions with applications to relativistic Quantum Field Theory.* Danske Videnskabernes Selskab Matematisk-Fysiske Meddelelser 31, 1-41 (1957)

[WiGa64] A. S. Wightman, L.Gårding : *Fields as operator-valued distributions in relativistic quantum theory.* Arkiv för Fysik, 28, 129 - 184 (1964)

[Wil75] K. Wilson : *The renormalisation group: critical phenomena and the Kondo problem.* Rev. Mod. Phys. 47, 773 (1975)

[Y11] K.Yeats: *Rearranging Dyson-Schwinger equations*, AMS, Vol. 211, 995, arXiv: 'Growth estimates for Dyson-Schwinger equations', hep-th/08102249 (2011) (Different sign conventions!)

[Y13] K.Yeats, *Some combinatorial interpretations in perturbative quantum field theory,* arXiv: math-th/1302.0080

[YaFe50] C.N. Yang, D. Feldman: *The S-Matrix in the Heisenberg Representation.* Phys. Rev. 79, 972 (1950)

[Yng77] J. Yngvason : *Remarks on the reconstruction theorem for field theories with indefinite metric.* Rep. Math. Phys. 12, 57 -64 (1977)

[Zeid06] E.Zeidler: *Quantum Field Theory I: Basics in Mathematics and Physics.* Springer (2006)

Selbständigkeitserklärung

Hiermit erkläre ich, Lutz Klaczynski, dass ich die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt habe. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt.

Lutz Klaczynski